

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2002, 10:44:50 ; Search time 65.87 Seconds
(without alignments)
625.061 Million cell updates/sec

Title: US-09-733-756-2
Perfect score: 1301
Sequence: 1 MAAAPLLLLLLVPVPLPL.....REHLPLAVGIAFFLMTFLAS 238

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 562222 seqs, 172994929 residues 562222
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_19:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archheap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	727	55.9	219	13 Q90YC5	Q90yc5 brachydanio
2	586	45.0	118	11 Q923G4	Q923g4 rattus norv
3	424.5	32.6	202	13 Q98TZ1	Q98tz1 gallus gall
4	406	31.2	88	13 Q9PTD0	Q9ptd0 ctenophorus
5	406	31.2	206	11 Q9CZS8	Q9czs8 mus musculu
6	373	28.7	229	13 Q93431	Q93431 brachydanio
7	367	28.2	205	11 Q9D7K8	Q9d7k8 mus musculu
8	338	26.0	102	11 Q9WUE7	Q9wue7 rattus norv
9	268.5	20.6	93	13 Q9PTD1	Q9ptd1 ctenophorus
10	247.5	19.0	80	13 Q42304	O42304 brachydanio
11	208	16.0	42	11 Q9CRL3	Q9crl3 mus musculu
12	178	13.7	70	13 Q90ZG7	Q90zg7 brachydanio
13	166.5	12.8	331	13 Q90Z31	Q90z31 brachydanio
14	164.5	12.6	327	13 Q9PT69	Q9pt69 xenopus lae
15	162.5	12.5	333	13 Q9PUJ4	Q9puj4 gallus gall
16	145	11.1	652	5 Q9V4E1	Q9v4e1 drosophila

17	143.5	11.0	348	5	O44516	O44516 caenorhabdi
18	138.5	10.6	279	5	Q9U474	Q9u474 caenorhabdi
19	138	10.6	334	13	Q90Z32	Q90z32 brachydanio
20	124.5	9.6	341	13	Q90Z33	Q90z33 brachydanio
21	100	7.7	377	3	O74823	O74823 schizosacch
22	96	7.4	1199	5	Q9V9Y3	Q9v9y3 drosophila
23	92	7.1	855	15	Q9YKQ0	Q9ykq0 human immun
24	88.5	6.8	441	4	Q9NQ99	Q9nq99 homo sapien
25	88.5	6.8	709	4	Q9NRN2	Q9nrn2 homo sapien
26	88.5	6.8	894	4	Q9UMZ9	Q9umz9 homo sapien
27	87.5	6.7	765	11	Q91ZV1	Q91zv1 mus musculu
28	87.5	6.7	765	11	Q91V98	Q91v98 mus musculu
29	87.5	6.7	894	4	Q9BZH4	Q9bzh4 homo sapien
30	87	6.7	817	5	Q93560	Q93560 caenorhabdi
31	86	6.6	438	11	Q9R262	Q9r262 mus musculu
32	86	6.6	442	11	Q9Z2L7	Q9z2l7 mus musculu
33	85.5	6.6	709	5	O16783	O16783 caenorhabdi
34	84	6.5	855	15	Q9YKQ4	Q9ykq4 human immun
35	84	6.5	1027	5	Q17657	Q17657 caenorhabdi
36	83.5	6.4	457	3	Q05672	Q05672 saccharomyc
37	83.5	6.4	473	5	Q961A1	Q961a1 drosophila
38	83.5	6.4	488	3	Q07619	Q07619 saccharomyc
39	83.5	6.4	502	5	Q9W0V9	Q9w0v9 drosophila
40	83.5	6.4	561	5	Q9V9N4	Q9v9n4 drosophila
41	83.5	6.4	910	11	Q9JIL3	Q9jil3 rattus norv
42	83	6.4	1342	5	Q9GPP6	Q9gpp6 drosophila
43	83	6.4	2206	12	Q9DIF1	Q9dif1 human polio
44	82.5	6.3	304	4	Q9HBC3	Q9hbc3 homo sapien
45	82.5	6.3	463	4	Q13648	Q13648 homo sapien

ALIGNMENTS

RESULT 1
Q90YC5 ID Q90YC5 PRELIMINARY; PRT; 219 AA.
AC Q90YC5;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE EPHRIN-A3.
GN EPHRIN-A3.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21412237; PubMed=11520665;
RA Hirate Y., Mieda M., Harada T., Yamasu K., Okamoto H.;
RT "Identification of ephrin-A3 and novel genes specific to the midbrain-
RT MHB in embryonic zebrafish by ordered differential display.";
RL Mech. Dev. 107:83-96(2001).
DR EMBL; AB051678; BAB55891.1; -.
SQ SEQUENCE 219 AA; 25146 MW; 7191927E03F8EA01 CRC64;

Query Match 55.9%; Score 727; DB 13; Length 219;
Best Local Similarity 61.9%; Pred. No. 1.4e-67;
Matches 140; Conservative 25; Mismatches 43; Indels 18; Gaps 3;

QY	1	MAAAPLLLLLLVPVPLLP	LAQPGGALGNRHAVYWNSSNOHLRREGYTVQVNVNDYLD	60
Db	1	MALVALCLFLTLTCTNFALVTAA	-----RHAVHWNSSNILLRKEGYTLQVNVNDYLD	53
QY	61	IYCPHYNSSGVGPGAGPGGGA	EQYVLYMVSRNGYRTCNASQCFKRWECNRPHAPHSPI	120
Db	54	IYCPHYNSS-----	QRGIAEQYVLYMVSYGYRTCDPOLGFKRWECNRPHAPAPI	104
QY	121	KFSEKFORYSAFSLGYEFHAG	HEYIYSTPTNHLHWKCLRMKVFVCCASTSHSGEKPVPT	180

Db 105 KFSEKQRYSAFSLGYEFHVGQYIYSTPTHHHGRSCLRLRVYVCCSTASDDEPQPT 164

QY 181 LPOFTMGPNVKINVLDFEGENPQVPKLEKISGTSPPKREHLPLAV 226

Db 165 EPDYTLRPNIKIDDDY--DNPEVPKLEKISGSSPSRDRLLTV 208

RESULT 2

Q923G4

ID Q923G4 PRELIMINARY; PRT; 118 AA.

AC Q923G4;

DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE EPHRIN A3 (FRAGMENT).

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RA Munoz J.J., Alonso-C L.M., Sacedon R., Crompton T., Vicente A.,

RA Jimenez E., Varas A., Zapata A.G.;

RT "Expression and function of the Eph A receptors and their ligands

RT ephrins A in the rat thymus."

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY045577; AAK92219.1; -.

FT NON_TER 1

FT NON_TER 118

SQ SEQUENCE 118 AA; 13470 MW; FF0277F079783A46 CRC64;

Query Match 45.0%; Score 586; DB 11; Length 118;

Best Local Similarity 72.4%; Pred. No. 3e-53;

Matches 110; Conservative 3; Mismatches 5; Indels 34; Gaps 2;

QY 63 CPHYNSSGVGPGAGPGGGGAEQYVLYMVRNGYRTCNASQGFKRWEKNRPHAPHSPIKF 122

Db 1 CPHYNSS-----GPGGGAEQYVLYMVNLGYRTCNASQGSKRWEKNRQHSHSPIKF 52

QY 123 SEKQRYSAFSLGYEFHAGHEYIYSTPTHNLHWKCLRMKVFVCCASTSHSKEKVPVPTLP 182

Db 53 SEKQRYSAFSLGYEFHAGQYIYSTPTHNLHWKCLRMKVFVCCAS----- 99

QY 183 QFTMGPNVKINVLDFEGENPQVPKLEKSISG 214

Db 100 -----KDFEGENPQVPKLEKSISG 118

RESULT 3

Q98T21

ID Q98T21 PRELIMINARY; PRT; 202 AA.

AC Q98T21;

DT 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)

DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)

DE EPHRIN-A6 (FRAGMENT).

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RA Menzel P., Valencia F., Godement P., Dodelet V.C., Pasquale E.B.;

RT "Ephrin-A6, a new ligand for EphA receptors in the developing visual

RT system."

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF317286; AAK00944.1; -.

DR InterPro; IPR001799; Ephrin.

DR Pfam; PF00812; Ephrin; 1.

DR PRINTS; PR01347; EPHRIN.

DR ProDom; PD002533; Ephrin; 1.

DR PROSITE; PS01299; EPHRIN; 1.

FT NON_TER 1

SQ SEQUENCE 202 AA; 22624 MW; 91E2716FF91353F9 CRC64;

Query Match 32.6%; Score 424.5; DB 13; Length 202;

Best Local Similarity 45.4%; Pred. No. 3.8e-36;

Matches 84; Conservative 28; Mismatches 56; Indels 17; Gaps 4;

QY 6 LLLLLLLVPVPLLLAQQPGGALGNRHAVYWNSSNQHLRRREGYTVQVNVNDYLDIYCPH 65

Db 1 LLGLLLWA-----PLLWAPPPVGRRRHGVYWNSSNPRFLQDDYSIQVSIQVNDHLDIYCPH 55

QY 66 YNSSGVGPGAGPGGGGAEQYVLYMVRNGYRTCNASQG-FKRWEKNRPHAPHSPIKFSE 124

Db 56 YSA-----PTPWAESFTLFMVDEEGYRGCSSETPGAFKRWEKNRPFAPFVPVRFSE 105

QY 125 KFQRYSAFSLGYEFHAGHEYIYSTPTHNLHWKCLRMKVFVCCASTSHSKEKVPV-TLPQ 183

Db 106 KIQRFTPSLGFEEFPGETYYIISVPTPGSAGRCCLKLRVSVCCRASTPEPLTEVPNSQPR 165

QY 184 FTMGP 188

Db 166 GRGGP 170

RESULT 4

Q9PTD0

ID Q9PTD0 PRELIMINARY; PRT; 88 AA.

AC Q9PTD0;

DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

DE EPHRIN A3 (FRAGMENT).

OS Ctenophorus ornatus.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Amphibolurinae;

OC Ctenophorus.

OX NCBI_TaxID=95347;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;

RA Chen P.B., Rodger J., Dunlop S.A., Beazley L.D.;

RT "Ephrin homologs are expressed in the adult lizard visual system."

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;

RA Scarborough M.T.C., Rodger J., Dunlop S.A., Beazley L.D.;

RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF209777; AAF19444.1; -.

DR InterPro; IPR001799; Ephrin.

DR Pfam; PF00812; Ephrin; 1.

DR ProDom; PD002533; Ephrin; 1.

FT NON_TER 1

FT NON_TER 88

SQ SEQUENCE 88 AA; 10636 MW; C56FCD8B13F219E7 CRC64;

Query Match 31.2%; Score 406; DB 13; Length 88;

Best Local Similarity 75.0%; Pred. No. 1.1e-34;

Matches 72; Conservative 7; Mismatches 9; Indels 8; Gaps 1;

QY 35 YWNSSNQHLRRREGYTVQVNVNDYLDIYCPHYNSSGVGPGAGPGGGGAEQYVLYMVRN 94

Db 1 VFWNRSNPHLRREGYTVQVSVNDYLDIYCPHYNAS-----VPEGRVEQYILYMVNYE 52

QY 95 GYRTCNASQGFKRWEKNRPHAPHSPIKFSEKQRYIS 130

Db 53 GYRTCNISQGFKRWEKNRPHAPHSPIKFSEKQRFRT 88

RESULT 5


```
Q9CZS8
ID Q9CZS8 PRELIMINARY; PRT; 206 AA.
AC Q9CZS8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 10 DAYS EMBRYO CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:2610529M21, FULL INSERT SEQUENCE.
GN EFNA4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK012195; BAB28092.1; -.
DR MGD; MGI:106643; Efn4.
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin; 1.
DR PRINTS; PR01347; EPHRIN.
DR PRODOM; PD002533; Ephrin; 1.
DR PROSITE; PS01299; EPHRIN; 1.
SQ SEQUENCE 206 AA; 22859 MW; 675E32971D1C6EBC CRC64;
```

```
Query Match 31.2%; Score 406; DB 11; Length 206;
Best Local Similarity 46.9%; Pred. No. 3.e-34;
Matches 82; Conservative 19; Mismatches 62; Indels 12; Gaps 3;

QY 8 LLLLLVPVPLPLLAQGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLCYPHYN 67
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3 LLPLLRVTWLWALLGSLRPGCSSLRHPIYWNSSNPRLLRGDAVVELGFNDYLDIFCPHYE 62
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 68 SSGVGPGAGPGPGGGAQYVLYMVSRNGYRTCA--SQGFKRWEKNRPHAPHSPIKFSEK 125
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 63 S-----PGPPEGTFALYIVDWSGYEACTAEGANSFORWNCSMPFAPFSVRFSEK 114
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 126 FQYSAFSLGYEFHAGHEYHYISTPTNHLHWKCLRKMVFVCC--ASTSHSGKPV 178
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 115 IQRYTPFLGFEFLPGETYYISVPTPESPGRCRLQLQVSVCKESGSSHAHPV 169
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 6
O93431
ID O93431 PRELIMINARY; PRT; 229 AA.
AC O93431;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE EPHRIN A-L1.
OS Brachydanio rerio (zebrafish) (zebra danio).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Durbin L., Brennan C.H., Shiomi K., Cooke J.;
RT "Eph signalling is required for segmentation and differentiation of
   the somites.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ006838; CAA07264.1; -.
DR InterPro; IPR001799; Ephrin.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00812; Ephrin; 1.
DR PRINTS; PR01347; EPHRIN.
DR PRODOM; PD002533; Ephrin; 1.
DR PROSITE; PS01299; EPHRIN; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN1.
SQ SEQUENCE 229 AA; 26115 MW; 8684462F67AF6F5C CRC64;
```

```
Query Match 28.7%; Score 373; DB 13; Length 229;
Best Local Similarity 40.0%; Pred. No. 1e-30;
Matches 84; Conservative 25; Mismatches 59; Indels 42; Gaps 6;

QY 28 ALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLCYPHYNSSGVGPGAGPGGGAEQYV 87
   |  |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 18 ASAEHRSVYWNSTNANFLWDDYTVDRINDYLDIICPHY-----AHGEIASQEAERYV 70
   |  |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 88 LYMVSRRNGYRTCA--SQGFKRWEKNRPHAPHSPIKFSEKFORYSAFSLGYEFHAGHEYV 146
   |||||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 71 LYMVELEDYENCKPHSFDQLRWCSRPFAPHAPEKFSEKFORFTPTLKGKFRQGESYV 130
   |||||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 147 ISTPTNHLHWKCLRKMVFVCCASTSHSGEKVPVTLPQFTMGPNVKINLVLEDFEGE----- 201
   |||||  |||  |||||  |||  |||  |||  |||  |||  |||  |||  |||
Db 131 ISKPLHHHGOECLRLKVDVVGPHGSKNKK-----MVEKVEIEGKMAAGG 176
   |||||  |||  |||||  |||  |||  |||  |||  |||  |||  |||  |||

QY 202 --NP-----QVPKLEKSI--SGTS 216
   |||||  |||  |||||  |||  |||  |||  |||  |||  |||  |||  |||
Db 177 VHNPSNRLPADDPIAMIPVQVRSVSGSGVS 206
   |||||  |||  |||||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 7
Q9D7K8
ID Q9D7K8 PRELIMINARY; PRT; 205 AA.
AC Q9D7K8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ADULT MALE TONGUE CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:2310004J15, FULL INSERT SEQUENCE.
GN EFNA1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
```

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK009144; BAB26102.1; -.
DR MGD; MGI:103236; Efnal.
DR InterPro; IPR001799; Ephrin.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00812; Ephrin; 1.
DR PRINTS; PR01347; EPHRIN.
DR ProDom; PD002533; Ephrin; 1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
SQ SEQUENCE 205 AA; 23772 MW; E37E55767459A4EC CRC64;

Query Match 28.2%; Score 367; DB 11; Length 205;
Best Local Similarity 38.3%; Pred. No. 3.7e-30;
Matches 85; Conservative 23; Mismatches 88; Indels 26; Gaps 6;

QY 19 PLLAQQPGGALGNRHAVYWNSSNQHLRRREGYTVQVNVNDYDIYCPHYNSSGVGPGAGPG 78
Db 7 PLLGLCCSLAAADRHIVFWNSSNPKFREEDYTVHVQLNDYDIICPHYEDDSVADAA--- 63

QY 79 PGCGAEQVLYMVSNGYRTCN-ASQGFKRWCNRPAPHSPHPIKFSEKQRYSAFSLGYE 137
Db 64 ---MERYTLYMVEHQEVVACQPSQKQVWNCNRPNSAKHGPVKLQVQFRTFFILGKE 119

QY 138 FHAGHEYIYSTPTNHLHWKCLRMKVFVCCASTSHSGEKPVPPTLPQFTMGPNVKINVLED 197
Db 120 FKEGHSYIYISKPIYHQESQCLKLVTV-----NG-----KITHNPQAHVNPQEK 164

QY 198 -FEGENPOVPKLEKSISGTSKREHLPLAVGIAFFLMTFLAS 238
Db 165 RLQADDPEVQVLSIGYSAAPRL--FPLVWAVLLPLLLQLQS 204

RESULT 8
Q9WUE7 PRELIMINARY; PRT; 102 AA.
AC Q9WUE7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE EPHRIN A-2 (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Matsunaga T., Davis J.G., Greene M.I.;
RT "Cellular and subcellular compartmentalization of adult peripheral
RT vestibular system by distinctive and overlapping expression of Eph
RT receptors and ephrins.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AF131912; AAD33515.1; -.
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin; 1.
DR ProDom; PD002533; Ephrin; 1.
DR PROSITE; PS01299; EPHRIN; 1.
FT NON_TER 1
FT NON_TER 102
SQ SEQUENCE 102 AA; 11838 MW; A497302F7FD7364B CRC64;

Query Match 26.0%; Score 338; DB 11; Length 102;
Best Local Similarity 58.9%; Pred. No. 1.6e-27;
Matches 63; Conservative 13; Mismatches 23; Indels 8; Gaps 2;

QY 49 YTVQVNVNDYDIYCPHYNSSGVGPGAGPGGGAQYVLYMVSNGYRTCNASQ-GFKR 107

Db 1 YTVESINDYLDIYCPHY-----GAPLPPAERMERYILYMVNGEGHASCDCRQGFKR 53
QY 108 WECNRPAPHSPHPIKFSEKQRYSAFSLGYEFHAGHEYIYSTPTNHL 154
Db 54 WECNRPAAAGGGLKFKSEKQLFTFSLGFEPFGHEYYIISATPPNL 100

RESULT 9
Q9PTD1 PRELIMINARY; PRT; 93 AA.
AC Q9PTD1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE EPHRIN A2 (FRAGMENT).
OS Ctenophorus ornatus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Amphibolurinae;
OC Ctenophorus.
OX NCBI_TaxID=95347;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Scarborough M.T.C., Rodger J., Dunlop S.A., Beazley L.D.;
RT "Ephrin homologs are expressed in the adult lizard visual system.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AF209776; AAF19443.1; -.
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin; 1.
DR ProDom; PD002533; Ephrin; 1.
FT NON_TER 1
FT NON_TER 93
SQ SEQUENCE 93 AA; 11323 MW; BD561F18D34C0F28 CRC64;

Query Match 20.6%; Score 268.5; DB 13; Length 93;
Best Local Similarity 49.5%; Pred. No. 2.4e-20;
Matches 48; Conservative 18; Mismatches 22; Indels 9; Gaps 2;

QY 35 VYWNSSNQHLRRREGYTVQVNVNDYDIYCPHYNSSGVGPGAGPGGGAQYVLYMVSRL 94
Db 1 VFCNRSNPRFWQGEYTVAVSINDYLDVYCPYYES-----PQHSRMERYILFMVNH 52

QY 95 GYRTC-NASQGFKRWCNRPAPHSPHPIKFSEKQRY 130
Db 53 GYLTCHEHRMRGFKRWCNRPQSPDGLRFSEKQFRT 89

RESULT 10
O42304 PRELIMINARY; PRT; 80 AA.
AC O42304;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE EPHRIN-A5 (FRAGMENT).
GN EFNA5A.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA MacDonald R., Scholes J., Strahle U., Brennan C., Holder N., Brand M.,
RA Wilson S.;
RT "The Pax protein Noi protein is required for commissural axon pathway
RT formation in the rostral forebrain.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL: Y12928; CAA73391.1; -.
DR ZFIN; ZDB-GENE-001128-1; efna5a.
DR InterPro; IPR001799; Ephrin.


```
Db 12 ILLIFLVLLGITAT-----NMEPIYWNLSNKRFSDDKGYVLVYPQIGDRDLICPSSD 64
QY 68 SSGVGPGAGPGGGGA--EQYVLYMVSR-----NGYRTCNASQGFKRWECNR 112
Db 65 P-----PGPRADYIYKLYLVSSREQADRCEVTGAPNLLLTCD----- 104
QY 113 PHAPHSPIKFSEKFORYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVVFVCCASTSH 172
Db 105 --KPNSDMRFTIKFQYSPNLWGHEFKTNHDFIIAT-SDGTRQGLSMRGGVCAT---- 157
QY 173 SGEKPVPTLPQFTMGPNVK-----INVLEDFEGE--NPQVPKLEKSISGTSPKREH 221
Db 158 QGMKVVLKVGQSPYGLPAKSPKPDASGRINPNPNTGNSTHPQIPP-----RSGSGGENP 212
QY 222 LP-----LAVGIAFFLM 233
Db 213 LPASNIAVIAGAAGGSAFLLL 233

RESULT 14
Q9PT69 PRELIMINARY; PRT; 327 AA.
AC Q9PT69;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE EPHRIN-B3 PRECURSOR.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYO;
RX MEDLINE=20099673; PubMed=106333856;
RA Helbling P.M., Saulnier D.M.E., Robinson V., Christiansen J.H.,
RA Wilkinson D.G., Brandli A.W.;
RT "Comparative analysis of embryonic gene expression defines potential
RT interaction sites for Xenopus EphB4 receptors with ephrin-B ligands.";
RL Dev. Dyn. 216:361-373(1999).
DR EMBL; AJ236866; CAB65511.1; -.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin; 1.
DR ProDom; PD002533; Ephrin; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS01299; EPHRIN; 1.
KW Signal.
FT SIGNAL.
FT SIGNAL.
SQ SEQUENCE 327 AA; 35913 MW; 4BB0FA39D4C22DCD CRC64;

Query Match 12.6%; Score 164.5; DB 13; Length 327;
Best Local Similarity 26.6%; Pred. No. 8e-09;
Matches 63; Conservative 31; Mismatches 78; Indels 65; Gaps 13;

QY 35 VYWNSSNQHLR-REGYTVQVNVNDYLDIYCPHYNSSGVGPGAGPGGGAEQYVLYMV-- 91
Db 30 IYWNSSNKRFEDETEGYVLVYPQIGDRDLICRSEPO-----GPFSSSPYEEKLYLVGT 83
QY 92 -----SRNGYRTCNASQGFKRWECNRPHAPHSPIKFSEKFORYSAFSLGYEFHAGHEYIY 146
Db 84 KEEMSSCSILRTPNL-----LLTCDR---PSQDLRFTIKFQEFSPNLWGHEFQSORDYI 135
QY 147 ISTPTNHL-----HWKCLRMKVVFVCCASTSHSGEKPVPPTLPQFT-----MGPNV 190
Db 136 IATSDGTMGDIETLQGGVCETGKMKVTLKV---GQSPNATP-PRRPSSAGKDSGISPSV 191
QY 191 KINVLEFEGENPQVPKLEKSISGTSPKR-----EHLPL----AVGIAFFLMTF 235
Db 192 -----PNPDIPNVGET-SGNATKTGNGPLPISHVPLVAGAAGGALLLVF 237
```

```
RESULT 15
Q9PUJ4 PRELIMINARY; PRT; 333 AA.
AC Q9PUJ4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE EPHRIN-B2 PRECURSOR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Menzel P., Pasquale E.B.;
RT "Coding sequence of chicken ephrin-B2.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AF180729; AAD53948.1; -.
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin; 1.
DR ProDom; PD002533; Ephrin; 1.
DR PROSITE; PS01299; EPHRIN; 1.
KW Signal.
FT SIGNAL.
FT CHAIN.
FT CHAIN.
SQ SEQUENCE 333 AA; 36761 MW; 4C28E8CB211B7783 CRC64;
```

```
Query Match 12.5%; Score 162.5; DB 13; Length 333;
Best Local Similarity 27.1%; Pred. No. 1.3e-08;
Matches 61; Conservative 29; Mismatches 92; Indels 43; Gaps 9;

QY 35 VYWNSSN-QHLRREGYTVQVNVNDYLDIYCPHYNSSGVGPGAGPGGGAEQYVLYMVSR 93
Db 35 IYWNSSNPKFLPGQGLVLYPQIGDKLDIICPKVDSKTVGQ-----YEYKVMVDK 85
QY 94 NGYRTCNASQGFKRWECNRP-----HAPHSPIKFSEKFORYSAFSLGYEFHAGHEYIYST 149
Db 86 DQADSCAIRKD-----NTPLLNCAPKDDQDVKFTIKFQEFSPNLWGLEFQKNKDYIYST 139
QY 150 PTHNLH-----WKCLRMKVVFVCCASTSHSGEKPVPPT-----LPQFTMGPNVKINV-- 194
Db 140 SNGSLEGLNNQEGVCQTKMKILMKVGDQPNAGLPRSTDPTRPEQEAGTNGKSSTTS 199
QY 195 --LEDFEGENPQVPKL-EKSI-SGTSPKREHLPLAVGIAFFLMTFL 236
Db 200 PFVKDHSGSTDGSKAGHSSIIIGS-----EVALFAGIASGCIIFI 239
```

Search completed: July 13, 2002, 10:52:33
Job time: 463 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: July 13, 2002, 10:45:35 ; Search time 24.34 Seconds
(without alignments)
378.605 Million cell updates/sec

Title: US-09-733-756-2
Perfect score: 1301
Sequence: 1 MAAAPLLLLLLVVPVLLPL.....REHLPLAVGIAFFLMTFLAS 238

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1301	100.0	238	1 EFA3_HUMAN	P52797 homo sapien
2	963	74.0	187	1 EFA3_MOUSE	O08545 mus musculus
3	437	33.6	209	1 EFA2_MOUSE	P52801 mus musculus
4	437	33.6	213	1 EFA2_HUMAN	O43921 homo sapien
5	426	32.7	195	1 EFA2_BRARE	P79727 brachydanio
6	421	32.4	200	1 EFA2_CHICK	P52802 gallus gall
7	415.5	31.9	228	1 EFA5_HUMAN	P52803 homo sapien
8	414.5	31.9	228	1 EFA5_MOUSE	O08543 mus musculus
9	414.5	31.9	228	1 EFA5_RAT	P97605 rattus norv
10	410	31.5	206	1 EFA4_MOUSE	O08542 mus musculus
11	404.5	31.1	228	1 EFA5_CHICK	P52804 gallus gall
12	397	30.5	228	1 EFA5_BRARE	P79728 brachydanio
13	390	30.0	205	1 EFA1_HUMAN	P20827 homo sapien
14	375.5	28.9	216	1 EFA1_XENLA	P52794 xenopus lae
15	374	28.7	205	1 EFA1_MOUSE	P52793 mus musculus
16	364.5	28.0	201	1 EFA4_HUMAN	P52798 homo sapien
17	362	27.8	205	1 EFA1_RAT	P97553 rattus norv
18	179	13.8	340	1 EFB3_HUMAN	O15768 homo sapien
19	177	13.6	340	1 EFB3_MOUSE	O35393 mus musculus
20	171.5	13.2	333	1 EFB2_HUMAN	P52799 homo sapien
21	171.5	13.2	336	1 EFB2_MOUSE	P52800 mus musculus
22	148.5	11.4	346	1 EFB1_HUMAN	P98172 homo sapien
23	147	11.3	332	1 EFB2_BRARE	O73874 brachydanio
24	139.5	10.7	345	1 EFB1_MOUSE	P52795 mus musculus
25	134	10.3	327	1 EFB1_XENLA	O13097 xenopus lae
26	132.5	10.2	345	1 EFB1_RAT	O13097 rattus norv
27	122.5	9.4	334	1 EFB1_CHICK	P52796 rattus norv
28	91	7.0	2206	1 POLG_POL32	O73612 gallus gall
29	88.5	6.8	894	1 ILF3_HUMAN	P06209 poliovirus
30	85	6.5	400	1 ILF3_MOUSE	Q12906 h interleuk
31	83.5	6.4	910	1 CASP_MOUSE	Q9cyd3 mus musculus
32	82.5	6.3	432	1 ENV2_RAT	O9jil3 rattus norv
33	82.5	6.3	466	1 KRUP_DROME	P11370 mus musculus
					P07247 drosophila

ALIGNMENTS

RESULT 1
EFA3_HUMAN

ID EFA3_HUMAN STANDARD; PRT; 238 AA.

AC P52797;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ephrin-A3 precursor (EPH-related receptor tyrosine kinase ligand 3)

DE (LERK-3) (EHK1 ligand) (EHK1-L).

GN EFN3 OR EPLG3 OR LERK3 OR EFL-2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95140419; PubMed=7838529;

RA Kozlosky C.J., Maraskovsky E., McGrew J.T., Vandenbos T.,

RA Teepe M., Lyman S.D., Srinivasan S., Fletcher F.A., Gayle R.B. III,

RA Cerretti D.P., Beckmann M.P.;

RT "Ligands for the receptor tyrosine kinases hek and elk: isolation of

RT cDNAs encoding a family of proteins.";

RL Oncogene 10:299-306(1995).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=95063919; PubMed=7973638;

RA Davis S., Gale N.W., Aldrich T.H., Maisonnier P.C., Lhotak V.,

RA Pawson T., Goldfarb M., Yancopoulos G.D.;

RT "Ligands for EPH-related receptor tyrosine kinases that require

RT membrane attachment or clustering for activity.";

RL Science 266:816-819(1994).

CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, SKELETAL MUSCLE, SPLEEN,

CC THYMUS, PROSTATE, TESTIS, OVARY, SMALL INTESTINE, AND PERIPHERAL

CC BLOOD LEUKOCYTES.

CC -!- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U14187; AAC50078.1; -.

EMBL; L37360; AAA52368.1; -.

MIM; 601381; -.

InterPro; IPR001799; Ephrin.

Pfam; PF00812; Ephrin; 1.

PRINTS; PR01347; EPHRIN.

ProDom; PD002533; Ephrin; 1.

PROSITE; PS01299; EPHRIN; 1.

Glycoprotein; GPI-anchor; Signal.

SIGNAL 1 22 POTENTIAL.

CHAIN 23 238 EPHRIN-A3.


```
EFA2_BRARE
ID EFA2_BRARE STANDARD; PRT; 195 AA.
AC P79727;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ephrin-A2 precursor (EPH-related receptor tyrosine kinase ligand 6)
DE (LERK-6) (ELF-1) (ZFEPHL3).
GN EFNA2 OR EPLG6 OR LERK6.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=97195707; PubMed=9043080;
RA Brennan C., Monschau B., Lindberg R., Guthrie B., Drescher U.,
RA Bonhoeffer F., Holder N.;
RT "Two Eph receptor tyrosine kinase ligands control axon growth and may
RT be involved in the creation of the retinotectal map in the
RT zebrafish.";
RL Development 124:655-664(1997).
CC -!- FUNCTION: CONTROL AXON GROWTH AND MAY BE INVOLVED IN THE CREATION
CC OF THE RETINO-TECTAL MAP.
CC
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC
CC -!- TISSUE SPECIFICITY: WIDESPREAD EXPRESSION IN THE EMBRYO.
CC
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE PRESUMPTIVE MIDBRAIN OF
CC DEVELOPING EMBRYOS FROM THE SIX-SOMITE STAGE. BY 24 HOURS IT IS
CC EXPRESSED THROUGHOUT THE MIDBRAIN INCLUDING THE REGION OF THE
CC PRESUMPTIVE TECTUM. AT LATER STAGES IT IS EXPRESSED IN A GRADED
CC FASHION THROUGHOUT THE TECTUM.
CC
CC -!- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y09668; CAA70863.1; .
CC ZFIN; ZDB-GENE-990415-66; efna2.
CC InterPro; IPR001799; Ephrin.
CC Pfam; PF00812; Ephrin; 1.
CC PRINTS; PR01347; EPHRIN.
CC ProDom; PD002533; Ephrin; 1.
CC PROSITE; PS01299; EPHRIN; 1.
CC Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 16
FT CHAIN 17 195
FT CARBOHYD 32 32
FT SEQUENCE 195 AA; 22688 MW; 9EE284FEB61D0C42 CRC64;
Query Match 32.7%; Score 426; DB 1; Length 195;
Best Local Similarity 52.3%; Pred. No. 4.9e-33;
Matches 81; Conservative 24; Mismatches 38; Indels 12; Gaps 4;
QY 29 LGNRHAVYWNSSNQHLRREGYTVQVNVNDYLDIYCAPHYNSGSGVGPAGPGGGAEQYVL 88
: : ||||| : ||| : ||||| : ||| : ||||| : ||| : ||||| : ||| : |||||
Db 23 ISDRHAVYWNSSNSRFWQGEYTVAVSINDYLDVYCPYIES-----PQPHSRMERYIL 74
QY 89 YMVSRNGYRVC-NASQGFKRWEKNCNRPAPHSPKIFSEKQRYSAFSLGYEFHAGHEYYII 147
: || : || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 75 FMVNHGDLTCEHRMRGFKRWEKNCNRPQSDGPLRSEKFLFTFSLGFEPFGHEYYII 134
QY 148 STPTHNLHWK-CLRMKVFVCCASTSHSGEKPVPPTL 181
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 135 SSPHPNHAGKPKCLKKVYV--KPTSSGYESPEPFL 167
```

```
RESULT 6
EFA2_CHICK
ID EFA2_CHICK STANDARD; PRT; 200 AA.
AC P52802;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ephrin-A2 precursor (EPH-related receptor tyrosine kinase ligand 6)
DE (LERK-6) (ELF-1).
GN EFNA2 OR EPLG6 OR LERK6 OR ELF1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95360981; PubMed=7634327;
RA Cheng H.J., Nakamoto M., Bergemann A.D., Flanagan J.G.;
RA "Complementary gradients in expression and binding of ELF-1 and Mek4
RA in development of the topographic retinotectal projection map.";
RL Cell 82:371-381(1995).
CC -!- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA3, EPHA4 AND
CC EPHA5 (BY SIMILARITY).
CC
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC
CC -!- TISSUE SPECIFICITY: EXPRESSED IN A GRADIENT ACROSS THE TECTUM
CC BEING MORE STRONGLY EXPRESSED AT THE POSTERIOR POLE.
CC
CC -!- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L40932; AAC42229.1; .
CC InterPro; IPR001799; Ephrin.
CC Pfam; PF00812; Ephrin; 1.
CC PRINTS; PR01347; EPHRIN.
CC ProDom; PD002533; Ephrin; 1.
CC PROSITE; PS01299; EPHRIN; 1.
CC Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 22
FT CHAIN 23 200
FT CARBOHYD 36 36
FT CARBOHYD 161 161
FT CARBOHYD 175 175
FT SEQUENCE 200 AA; 23049 MW; 8FAB1AE5E45EED96 CRC64;
Query Match 32.4%; Score 421; DB 1; Length 200;
Best Local Similarity 47.5%; Pred. No. 1.5e-32;
Matches 86; Conservative 29; Mismatches 50; Indels 16; Gaps 5;
QY 3 AAPLLLLLLLVPVPLPLLAQPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLDIY 62
: || : || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 6 AAALLAAIVGVCV-----WSDDPGKVISDRYAVYWNRSNPRFRHGRDYTEVSINDYLDIY 60
QY 63 CPHYNSSGVGPAGPGGGAEQYVLYMVSRNGYRVCNASQ-GFKRWEKNCNRPAPHSPK 121
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 61 CPHYEE-----PLPAERMERYVLYMVNYEGHSCDHRQKGFKRWEKNCNRPDPSGGLK 112
QY 122 FSEKQRYSAFSLGYEFHAGHEYYIIS-TPTHNLHWKCLRMKVFVCCASTS-HSGEKPVP 179
: |||| : |||| : |||| : |||| : |||| : |||| : |||| : |||| : |||| : ||||
Db 113 FSEKQLFTFSLGFEPFGHEYYIISASPLNVVDVDRPCLKKVYVRPTNDSLYESPEPIF 172
QY 180 T 180
```


Db 173 T 173

RESULT 7

EFA5_HUMAN STANDARD; PRT: 228 AA.

ID AC P52803;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Ephrin-A5 precursor (EPH-related receptor tyrosine kinase ligand 7)

DE (LERK-7) (AL-1).

GN EFNA5 OR EPLG7 OR LERK7.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95267434; PubMed=7748564;

RA Winslow J.W., Moran P., Valverde J., Shih A., Yuan J.Q., Wong S.C.,

RA Tsai S.P., Goddard A., Henzel W.J., Hefti F.;

RT "Cloning of AL-1, a ligand for an Eph-related tyrosine kinase

RT receptor involved in axon bundle formation.";

RL Neuron 14:973-981(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=97392664; PubMed=9245480;

RA Kozlosky C.J., Vandenbos T., Park L.S., Cerretti D.P., Carpenter M.K.;

RT "LERK-7: a ligand of the Eph-related kinases is developmentally

RT regulated in the brain.";

RL Cytokine 9:540-549(1997).

RN [3]

RP FUNCTION.

RX MEDLINE=20069483; PubMed=10601038;

RA Davy A., Gale N.W., Murray E.W., Klinghoffer R.A., Soriano P.,

RA Feuerstein C., Robbins S.M.;

RT "Compartmentalized signaling by GPI-anchored ephrin-A5 requires the

RT Fyn tyrosine kinase to regulate cellular adhesion.";

RL Genes Dev. 13:3125-3135(1999).

CC -!- FUNCTION: MAY FUNCTION ACTIVELY TO STIMULATE AXON FASCICULATION.

CC INDUCES COMPARTMENTALIZED SIGNALING WITHIN A CAVEOLAE-LIKE

CC MEMBRANE MICRODOMAIN WHEN BOUND TO THE EXTRACELLULAR DOMAIN OF ITS

CC COGNATE RECEPTOR. THIS SIGNALING EVENT REQUIRES THE ACTIVITY OF

CC THE FYN TYROSINE KINASE.

CC -!- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA2, EPHA3 AND

CC EPHB1.

CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR IT

CC IS COMPARTMENTALIZED IN DISCRETE CAVEOLAE-LIKE MEMBRANE

CC MICRODOMAINS.

CC -!- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U26403; AAB60377.1; -.

DR MIM; 601535; -.

DR InterPro; IPR001799; Ephrin.

DR Pfam; PF00812; Ephrin; 1.

DR PRINTS; PR01347; EPHRIN.

DR ProDom; PD002533; Ephrin; 1.

DR PROSITE; PS01299; EPHRIN; 1.

KW Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; Signal;

KW Polymorphism.

FT SIGNAL 1 20 POTENTIAL.

FT CHAIN 21 228 EPHRIN-A5.

FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARIANT 55 55 N -> K (IN DBSNP:469062).

FT /FTId=VAR_012035.

SQ SEQUENCE 228 AA; 26297 MW; 6893B1CCACFF3F57 CRC64;

Query Match 31.9%; Score 415.5; DB 1; Length 228;

Best Local Similarity 39.5%; Pred. No. 5.7e-32;

Matches 98; Conservative 37; Mismatches 66; Indels 47; Gaps 11;

QY 8 LLLLLVPLPLLAQPGG-ALGNRHAVYWNSSNQHLRRGYTVQVNVNDYLDIYCPHY 66

Db :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |

6 MLTLVFLVLMVCVFSQDPGSKAVADRYAVYWNSSNPRFQGDYHIDVCINDYLDVFCPHY 65

QY 67 NSSGVGPGACGPGGAEQYVLYMVSRNGYRTC-NASQGFKRWCNRPHPHSPKPFSEK 125

Db :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |

66 EDS-----VPEDKTERYVLYMVNFDGYSACDHTSKGFKRWCNRPHPSPNGPLKPFSEK 117

QY 126 FQYSAFSLGYEFHAGHEYIYSTP-THNLHWKCLRMKVFV-----CCASTSHSKEKPVPT 180

Db :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |

118 FQLFTPFLGFEFRPGREYFYISSAIPDNGRRSCLKLKVFRVPTNSCMKTIGVHDR----- 173

QY 181 LPQFTMGPNVKINLVLEDFE-----GEN-PQVPKLEKSISGTSFKREHLPLAVGI 228

Db :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |

174 --VFDVNDKVE-NSLEPADDTVHESAEPSRGENAAQTPTI-----PSR-----LLAI 217

QY 229 AFFLMTFL 236

Db :| | :| |

218 LLFLAML 225

RESULT 8

EFA5_MOUSE STANDARD; PRT: 228 AA.

ID AC O08543; O08544;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Ephrin-A5 precursor (EPH-related receptor tyrosine kinase ligand 7)

DE (LERK-7) (AL-1).

GN EFNA5 OR EPLG7 OR LERK7 OR EPL7.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97060319; PubMed=8903354;

RA Flenniken A.M., Gale N.W., Yancopoulos G.D., Wilkinson D.G.;

RT "Distinct and overlapping expression patterns of ligands for

RT Eph-related receptor tyrosine kinases during mouse embryogenesis.";

RL Dev. Biol. 179:382-401(1996).

CC -!- FUNCTION: MAY FUNCTION ACTIVELY TO STIMULATE AXON FASCICULATION.

CC INDUCES COMPARTMENTALIZED SIGNALING WITHIN A CAVEOLAE-LIKE

CC MEMBRANE MICRODOMAIN WHEN BOUND TO THE EXTRACELLULAR DOMAIN OF ITS

CC COGNATE RECEPTOR. THIS SIGNALING EVENT REQUIRES THE ACTIVITY OF

CC THE FYN TYROSINE KINASE (BY SIMILARITY).

CC -!- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA2, EPHA3 AND

CC EPHB1 (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR IT

CC IS COMPARTMENTALIZED IN DISCRETE CAVEOLAE-LIKE MEMBRANE

CC MICRODOMAINS (BY SIMILARITY).

CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A

CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -!- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

```
DR EMBL; U90664; AAB50239.1; -.
DR EMBL; U90665; AAB50240.1; -.
DR MGI; 107444; Efnas5.
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin; 1.
DR PRINTS; PR01347; EPHRIN.
DR ProDom; PD002533; Ephrin; 1.
DR PROSITE; PS01299; EPHRIN; 1.
KW Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; Signal;
KW Alternative splicing.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 228 EPHRIN-A5.
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 163 189 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 228 AA; 26339 MW; 85439F5337420022 CRC64;

Query Match 31.9%; Score 414.5; DB 1; Length 228;
Best Local Similarity 39.5%; Pred. No. 7.1e-32;
Matches 98; Conservative 36; Mismatches 67; Indels 47; Gaps 11;

QY 8 LLLLLVPVPLLLAQQPGG-ALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLDIYCPHY 66
Db :| | | | : : : | | : : | | | | | | : | : | : | | | | | |
6 MLTLFLVLWMCVFSQDPGSKVADRYAVYWNSSNPRFQGDYHIDVCINDYLDVFCPHY 65

QY 67 NSSGVGPGAGPGGGAEQYVLYMVSNGYRTC-NASQGFKRWECPNRPAPHSPKIFSEK 125
Db | : | | | | : | | : | | | | | | | | : | : | | | | | |
66 EDS-----VPEDKTERYVLYMVNFDGYSACDHTSKGFKRWECPNRPAPHSPKIFSEK 117

QY 126 FQYSAFSLGYEFHAGHEYIISTP-THNLHWKCLRMKFV-----CCASTSHSGEKPVP 180
Db || : | | | | | | | | | | | | | | | | | | | | | | | | | |
118 FQLTPFSLGFEPFGREYFISSAIPDNGRRSCLKLVFVRPTNSCMKTIGVHDR---- 173

QY 181 LPQFTMGPNVKINVLEDFE-----GEN-PQVPKLEKSISGTSKREHLPLAVGI 228
Db | : | : | | | | : | | | | | | | | | | | | | | | | |
174 --VFDVNDKVE-NSLEPADDTVHESAEPSPRGENAAQTPTI-----PSR-----LLAI 217

QY 229 AFFLMTFL 236
Db || : |
218 LLFLLAML 225

RESULT 9
EFA5_RAT STANDARD; PRT; 228 AA.
AC P97605;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ephrin-A5 precursor (EPH-related receptor tyrosine kinase ligand 7)
DE (LERK-7) (AL-1).
GN EFNA5 OR EPLG7 OR LERK7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=95267434; PubMed=7748564;
RA Winslow J.W., Moran P., Valverde J., Shih A., Yuan J.Q., Wong S.C.,
RA Tsai S.P., Goddard A., Henzel W.J., Hefti F., Beck K.D., Caras I.W.;
RT "Cloning of AL-1, a ligand for an Eph-related tyrosine kinase
RT receptor involved in axon bundle formation.";
RL Neuron 14:973-981(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Li Y.Y., McTiernan C.F., Feldman A.M.;
RT "rLERK7, rat ligand for Eph-related receptor tyrosine kinase.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY FUNCTION ACTIVELY TO STIMULATE AXON FASCICULATION.
CC INDUCES COMPARTMENTALIZED SIGNALING WITHIN A CAVEOLAE-LIKE
```

```
CC MEMBRANE MICRODOMAIN WHEN BOUND TO THE EXTRACELLULAR DOMAIN OF ITS
CC COGNATE RECEPTOR. THIS SIGNALING EVENT REQUIRES THE ACTIVITY OF
CC THE FYN TYROSINE KINASE (BY SIMILARITY).
CC -!- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA2, EPHA3 AND
CC EPHB1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR IT
CC IS COMPARTMENTALIZED IN DISCRETE CAVEOLAE-LIKE MEMBRANE
CC MICRODOMAINS (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, HEART, PLACENTA AND LUNG.
CC -!- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U69279; AAC05801.1; -.
CC InterPro; IPR001799; Ephrin.
CC Pfam; PF00812; Ephrin; 1.
CC PRINTS; PR01347; EPHRIN.
CC ProDom; PD002533; Ephrin; 1.
CC PROSITE; PS01299; EPHRIN; 1.
KW Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 228 EPHRIN-A5.
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 228 AA; 26358 MW; 855985532D580022 CRC64;

Query Match 31.9%; Score 414.5; DB 1; Length 228;
Best Local Similarity 39.5%; Pred. No. 7.1e-32;
Matches 98; Conservative 36; Mismatches 67; Indels 47; Gaps 11;

QY 8 LLLLLVPVPLLLAQQPGG-ALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLDIYCPHY 66
Db :| | | | : : : | | : : | | | | | | : | : | : | | | | | |
6 MLTLFLVLWMCVFSQDPGSKVADRYAVYWNSSNPRFQGDYHIDVCINDYLDVFCPHY 65

QY 67 NSSGVGPGAGPGGGAEQYVLYMVSNGYRTC-NASQGFKRWECPNRPAPHSPKIFSEK 125
Db | : | | | | : | | : | | | | | | | | : | : | | | | | |
66 EDS-----VPEDKTERYVLYMVNFDGYSACDHTSKGFKRWECPNRPAPHSPKIFSEK 117

QY 126 FQYSAFSLGYEFHAGHEYIISTP-THNLHWKCLRMKFV-----CCASTSHSGEKPVP 180
Db || : | | | | | | | | | | | | | | | | | | | | | | | | | |
118 FQLTPFSLGFEPFGREYFISSAIPDNGRRSCLKLVFVRPTNSCMKTIGVHDR---- 173

QY 181 LPQFTMGPNVKINVLEDFE-----GEN-PQVPKLEKSISGTSKREHLPLAVGI 228
Db | : | : | | | | : | | | | | | | | | | | | | | | | |
174 --VFDVNDKVE-NSLEPADDTVHESAEPSPRGENAAQTPTI-----PSR-----LLAI 217

QY 229 AFFLMTFL 236
Db || : |
218 LLFLLAML 225

RESULT 10
EFA4_MOUSE STANDARD; PRT; 206 AA.
AC O08542; O55218;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ephrin-A4 precursor (EPH-related receptor tyrosine kinase ligand 4)
DE (LERK-4).
GN EFNA4 OR EPLG4 OR LERK4 OR EPL4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
```

RX MEDLINE=97060319; PubMed=8903354;
RA Flenniken A.M., Gale N.W., Yancopoulos G.D., Wilkinson D.G.;
RT "Distinct and overlapping expression patterns of ligands for
RT Eph-related receptor tyrosine kinases during mouse embryogenesis";
RL Dev. Biol. 179:382-401(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=98126446; PubMed=9465306;
RA Cerretti D.P., Nelson N.;
RT "Characterization of the genes for mouse LERK-3/Ephrin-A3 (Epl3),
RT mouse LERK-4/Ephrin-A4 (Epl4), and human LERK-6/Ephrin-A2 (EPLG6):
RT conservation of intron/exon structure";
RL Genomics 47:131-135(1998).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
CC similarity).
CC -!- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U90663; AAC50238.1; -.
CC EMBL; U92890; AAC39962.1; -.
CC EMBL; U92889; AAC39962.1; JOINED.
CC MGD; MGI:106643; Efn4.
CC InterPro; IPR001799; Ephrin.
CC Pfam; PF00812; Ephrin; 1.
CC PRINTS; PR01347; EPHRIN.
CC ProDom; PD002533; Ephrin; 1.
CC PROSITE; PS01299; EPHRIN; 1.
CC Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 206 EPHRIN-A4.
FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 41 43 CELL ATTACHMENT SITE (POTENTIAL).
FT CONFLICT 1 4 MRL -> MLLRLGLIYPPTRPPAPPGLV (IN REF.
FT 1).
SQ SEQUENCE 206 AA; 22861 MW; 43501971DD1C6EA5 CRC64;

Query Match 31.5%; Score 410; DB 1; Length 206;
Best Local Similarity 47.4%; Pred. No. 1.7e-31;
Matches 83; Conservative 18; Mismatches 62; Indels 12; Gaps 3;

QY 8 LLLLLVPVPLLLAQQPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYDIYCPHYN 67
Db 3 LLPLLRITVLAALLGSRLPGCSSLRPIYWNSSNPRLLRGDAVVELGFDYDIFCPHYE 62

QY 68 SSGVGPAGPGGGAEQYVLYMVSRRNGYRTCA--SQGFKRWEENRPHAPHSPKFSK 125
Db 63 S-----PGPEGPETFALYMDWSGYEACTAEGANAFQWNCMPFAPFSPVRFSEK 114

QY 126 FQYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVFCV--ASTSHSGKPV 178
Db 115 IQRYTPPLGFEFLPGETYYIISVPTPESGRCLRLQVSVCKESGSSHSAHPV 169

RESULT 11
EFA5_CHICK
ID EFA5_CHICK STANDARD; PRT; 228 AA.
AC P52804;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ephrin-A5 precursor (EPH-related receptor tyrosine kinase ligand 7)
DE (LERK-7) (RAGS protein).
GN EFNA5 OR RAGS.

OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Posterior tectum;
RX MEDLINE=95360980; PubMed=7634326;
RA Drescher U., Kremoser C., Handwerker C., Loschinger J., Noda M.,
RA Bonhoeffer F.;
RT "In vitro guidance of retinal ganglion cell axons by RAGS, a 25 kDa
RT tectal protein related to ligands for Eph receptor tyrosine
RT kinases";
RL Cell 82:359-370(1995).
CC -!- FUNCTION: INDUCES GROWTH CONE COLLAPSE AND REPULSION OF RETINAL
CC -!- GANGLION CELL AXONS.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN A GRADED FASHION ACROSS THE
CC TECTUM BEING MORE STRONGLY EXPRESSED TOWARDS THE POSTERIOR POLE.
CC -!- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X90377; CAA62027.1; -.
CC InterPro; IPR001799; Ephrin.
CC Pfam; PF00812; Ephrin; 1.
CC PRINTS; PR01347; EPHRIN.
CC ProDom; PD002533; Ephrin; 1.
CC PROSITE; PS01299; EPHRIN; 1.
CC Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 228 EPHRIN-A5.
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 228 AA; 26206 MW; 56D8E4FBDECF18AD CRC64;

Query Match 31.1%; Score 404.5; DB 1; Length 228;
Best Local Similarity 49.1%; Pred. No. 6.2e-31;
Matches 79; Conservative 29; Mismatches 42; Indels 11; Gaps 4;

QY 8 LLLLLVPVPLLLAQQPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYDIYCPHY 66
Db 6 MLLLAVALWVCVRGQEPGRKAVADRYAVYWNSTNPRFQGDYHIDVCINDYLDVFCPHY 65

QY 67 NSSGVGPAGPGGGAEQYVLYMVSRRNGYRTCA--SQGFKRWEENRPHAPHSPKFSK 125
Db 66 EDS-----VPEDKTERYVLYMVSRRNGYRTCA--SQGFKRWEENRPHAPHSPKFSK 117

QY 126 FQYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVFCV 165
Db 118 FQLFTPFLGFEFRPGREYFYIISAPDNGRRSCLKLVFV 158

RESULT 12
EFA5_BRARE
ID EFA5_BRARE STANDARD; PRT; 228 AA.
AC P79728;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ephrin-A5 precursor (EPH-related receptor tyrosine kinase ligand 7)
DE (LERK-7) (AL-1) (ZFEPHL4).
GN EFNA5 OR EFNA5B OR AL1 OR EPLG7 OR LERK7.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

FT CONFLICT 204 204 S -> T (IN REF. 1).
SQ SEQUENCE 205 AA; 23802 MW; 5A8F3A6E2091E868 CRC64;

Query Match 28.7%; Score 374; DB 1; Length 205;
Best Local Similarity 38.7%; Pred. No. 4e-28;
Matches 86; Conservative 23; Mismatches 87; Indels 26; Gaps 6;

QY 19 PLLAQPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLDIYCPHYNSSGVGPGAGPG' 78
Db 7 PLLGLCCSLAAADRHIVFNSSNPKFREEDYTVHVQLNDYLDIICPHYEDDSVADAA--- 63

QY 79 PGGGAEQYVLYMVSRRNGYRTCN-ASOGFKRWECNRPHAPHSPKFKSEKFORYSAFSLGYE 137
Db 64 ----MERYTLYMVEHQEYVACQPQSKDQVRWNCNRPSAKHGPEKLSEKFORFTPPFILGKE 119

QY 138 FHAGHEYYYISTPTNHLHWKCLRMKVFCVCCASTSHSGEKPVPPTLPQFTMGPNVKINVLED 197
Db 120 FKEGHSYYIISKPIYHQESQCLKLKVTV-----NG-----KITHNPQAHVNPQEK 164

QY 198 -FEGENPQVPKLEKSISGTSPKREHLPLAVGTAFFLMTFLAS 238
Db 165 RLQADDPEVQVLHSHSIGYSAAPRL--FPLVWAVLLPLLLLOS 204

Search completed: July 13, 2002, 10:53:18
Job time: 463 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2002, 09:55:25 ; Search time 42.58 Seconds
(without alignments)
537.089 Million cell updates/sec

Title: US-09-733-756-2
Perfect score: 1301
Sequence: 1 MAAAPLLLLLLLLVPVPLPL.....REHPLAVGIAFFLMTFLAS 238

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1301	100.0	238	2 I38849	LERK-3 - human
2	437	33.6	209	2 A54984	ELF-1 protein precursor
3	434	33.4	213	2 JE0322	ephraim-A2 - human
4	415.5	31.9	228	2 I58170	LERK-7 precursor
5	404.5	31.1	228	2 A57084	repulsive axon guidance
6	390	30.0	205	2 A36377	B61 protein precursor
7	364.5	28.0	201	2 I38850	LERK-4 - human
8	171.5	13.2	333	2 I84743	hepatoma transmembrane
9	171.5	13.2	336	2 I49766	hepatoma transmembrane
10	148.5	11.4	346	2 S46993	elk ligand - human
11	143.5	11.0	462	2 T32645	hypothetical protein
12	139.5	10.7	345	2 I48780	Stral/Eplg2 protein
13	132.5	10.2	345	2 I58406	LERK-2 - rat
14	100	7.7	356	2 T40265	hypothetical zinc-finger
15	91	7.0	2206	1 GNNY27	genome polyprotein
16	88.5	6.8	1024	2 E90678	beta-D-galactosidase
17	88.5	6.8	1024	2 A85529	beta-D-galactosidase
18	87	6.7	817	2 T21336	hypothetical protein
19	85.5	6.6	709	2 T28712	hypothetical protein
20	84	6.5	1027	2 T19173	hypothetical protein
21	83.5	6.4	488	2 S67744	hypothetical protein
22	82.5	6.3	432	2 A25483	env polyprotein, r
23	82.5	6.3	466	1 TWFF	transcription factor
24	82	6.3	237	2 T19914	hypothetical protein
25	82	6.3	334	2 T16772	hypothetical protein
26	82	6.3	376	2 B84463	hypothetical protein
27	82	6.3	728	2 A48830	probable transcription
28	82	6.3	1144	2 A75132	hypothetical protein
29	81.5	6.3	1173	2 I50620	procr2 - chicken

30	81	6.2	457	2 T19109	hypothetical protein
31	81	6.2	1613	2 S39059	protein BRG1 - human
32	81	6.2	1647	2 S45252	SNF2beta protein - human
33	79.5	6.1	434	2 B72410	alkaline phosphatase
34	79	6.1	612	2 A54282	reversed polarity
35	79	6.1	859	1 VCLJMN	env polyprotein pr
36	78.5	6.0	202	2 T01605	phytoecyanin At2g44
37	78.5	6.0	498	2 B84789	probable protein w
38	78.5	6.0	1024	1 GBEC	beta-galactosidase
39	78.5	6.0	1291	2 T13389	hypothetical prote
40	78	6.0	321	2 I38238	transcription fact
41	78	6.0	328	1 S05426	chitinase (EC 3.2.
42	78	6.0	700	2 A32392	protein kinase C (
43	78	6.0	2206	1 GNNY4P	genome polyprotein
44	78	6.0	2206	2 S03822	genome polyprotein
45	77.5	6.0	570	2 H90370	thermopsin precurs

ALIGNMENTS

RESULT 1
I38849
LERK-3 - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-Sep-1999
C;Accession: I38849
R;Kozlosky, C.J.; Maraskovsky, E.; McGrew, J.T.; VandenBos, T.; Teepe, M.; Lyman, S.
Oncogene 10, 299-306, 1995
A;Title: Ligands for the receptor tyrosine kinases hek and elk: isolation of cDNAs e
A;Reference number: I38849; MUID:95140419
A;Accession: I38849
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-238 <RES>
A;Cross-references: EMBL:U14187; NID:g642832; PIDN:AAC50078.1; PID:g642833
C;Genetics:
A;Gene: GDB:EPLG3
A;Cross-references: GDB:438336; OMIM:601381
A;Map position: lq21-lq22
C;Superfamily: axon guidance signal protein

Query Match	100.0%	Score 1301;	DB 2;	Length 238;
Best Local Similarity	100.0%	Pred. No. 1.7e-114;		
Matches 238;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 1	MAAAPLLLLLLLLVPVPLPLLAQGGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60			
Db 1	MAAAPLLLLLLLLVPVPLPLLAQGGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60			
Qy 61	IYCPHYNSSGVGPGAGPGGGAEQYVLYMVSRNGYRTCNASQGFKECNRPHAPHSPI 120			
Db 61	IYCPHYNSSGVGPGAGPGGGAEQYVLYMVSRNGYRTCNASQGFKECNRPHAPHSPI 120			
Qy 121	KFSEKFORYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVFVCCASTSHSGEKPVT 180			
Db 121	KFSEKFORYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVFVCCASTSHSGEKPVT 180			
Qy 181	LPOFTMGPNVKINLVLEDFEGENPQVPKLEKISIGTSPKREHLPLAVGIAFFLMTFLAS 238			
Db 181	LPOFTMGPNVKINLVLEDFEGENPQVPKLEKISIGTSPKREHLPLAVGIAFFLMTFLAS 238			

RESULT 2
A54984
ELF-1 protein precursor - mouse
N;Alternate names: Cek7 ligand
C;Species: Mus musculus (house mouse)
C;Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 29-Sep-1999
C;Accession: A54984; A55873
R;Cheng, H.J.; Flanagan, J.G.
Cell 79, 157-168, 1994

R;Shao, H.; Lou, L.; Pandey, A.; Pasquale, E.B.; Dixit, V.M.

J. Biol. Chem. 269, 26606-26609, 1994

A;Title: cDNA cloning and characterization of a ligand for the Cek5 receptor protein-tyrosine kinase

A;Reference number: A55062; MUID:95014510

A;Accession: A55062

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-89, 'T', 91-345 <SHA>

A;Cross-references: GB:U12983; NID:g575928; PIDN:AAA53231.1; PID:g575929

C;Genetics:

A;Gene: EPLG2

Query Match 10.7%; Score 139.5; DB 2; Length 345;
Best Local Similarity 24.2%; Pred. No. 2.4e-05;
Matches 59; Conservative 27; Mismatches 87; Indels 71; Gaps 8;

QY 4 APLLLLLLVVPLLLPQAQGGALGNRRHAVYWNSSN-QHLRREGYTVQVNVNDYLDIY 62

Db 15 AMVVLTLCLRLATPLAK-----NLEPVSWSSLNPKFLSGKGLVIYPKIGDKLDII 63

QY 63 CPHYNSSGVGPGAGPGGGAQYVLYMV-----SRNGYRTCNASQGFKRWEC 110

Db 64 CPRAEA-----GRPYEYKLYLVRPEQAACSTVLDPNVLVTGN----- 102

QY 111 NRPHAPHSPIKFSKFORYSAFSLGYEFHAGHEYIYSTPTNLH-----WKCLRM 161

Db 103 ----KPHQEIRFTIKFQEFSPNYMGLFEPKHYDYYITSTNGSLEGLNREGGVCRTTM 158

QY 162 KVFVCCA-----STSHSGEKPVPPTLPQFTMGPNVKINLVLEDFEGENPQVPKLEK 210

Db 159 KIVMKVGQDPNAVTPPEQLTTSRPSKESDNTVKATQAPG--RGSQSDSGKHETVNOQEK 216

QY 211 SISO 214

Db 217 SGPG 220

RESULT 13

I58406

LERK-2 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 05-Nov-1999

C;Accession: I58406

R;Fletcher, F.A.; Carpenter, M.; Shilling, H.; Baum, P.; Ziegler, S.; Gimpel, S.; Hollin

Oncogene 9, 3241-3248, 1994

A;Title: LERK-2, a ligand for the receptor tyrosine kinase ELK, is evolutionarily conser

A;Reference number: I58406; MUID:95022634

A;Accession: I58406

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-345 <RES>

A;Cross-references: EMBL:U07560; NID:g563118; PIDN:AAA53092.1; PID:g563119

C;Genetics:

A;Gene: Eplg2

Query Match 10.2%; Score 132.5; DB 2; Length 345;
Best Local Similarity 23.8%; Pred. No. 0.00011;
Matches 58; Conservative 27; Mismatches 88; Indels 71; Gaps 8;

QY 4 APLLLLLLVVPLLLPQAQGGALGNRRHAVYWNSSN-QHLRREGYTVQVNVNDYLDIY 62

Db 15 AMVVLTLCLRLATPLAK-----NLEPVSWSSLNPKFLSGKGLVIYPKIGDKLDII 63

QY 63 CPHYNSSGVGPGAGPGGGAQYVLYMV-----SRNGYRTCNASQGFKRWEC 110

Db 64 CPRAEA-----GRPYEYKLYLVRPEQAACSTVLDPNVLVTGN----- 102

QY 111 NRPHAPHSPIKFSKFORYSAFSLGYEFHAGHEYIYSTPTNLH-----WKCLRM 161

Db 103 ----KPHQEIRFTIKFQEFSPNYMGLFEPKHYDYYITSTNGSLEGLNREGGVCRTTM 158

F;1028-1124/Product: core protein P2-5b #status predicted <P5B>
 F;1125-1453/Product: core protein P2-X #status predicted <P2X>
 F;1454-1540/Product: protein P3-1b #status predicted <P1B>
 F;1541-1562/Product: genome-linked protein VPg #status predicted <VPG>
 F;1563-1745/Product: proteinase #status predicted <PTS>
 F;1746-2206/Product: RNA-directed RNA polymerase #status predicted <RNS>
 F;1543/Binding site: phosphoryl-RNA (Tyr) (covalent) #status predicted

Query Match	7.0%;	Score 91;	DB 1;	Length 2206;
Best Local Similarity	20.5%;	Pred. No. 8.1;		
Matches 52;	Conservative	29;	Mismatches 109;	Indels 64; Gaps 9;

QY 26 GGALGNRHAV--YWNSSNQHLRREGYTVQVN-----VNDYLDIYCPHYNSSGVGPGAGPG 78

Db 21 GGSTINYTTTINYVKDSASNAASKODYSQDPFSKFTLEPLKDVLIKTAALNSPV----- 73

QY 79 PGGAEQYVLYMVSRNGYRT---CNASQGFKRW-----ECNRPHAPHSPIKFSEKFQ 127

Db 74 EACGYSRVLQLTLGNSTTTTQEAANSVWAYGRWPPEIFRDEANPVDQPTPEDVATSRFY 133

QY 128 RYSAFSLGYE-----FHAGHEYYYISTPTHNLHWKCLRMKVFVCCAST 170

Db 134 TLDTVMGKESRGWWWKLPDALRDMGLFGQNMYYHYLGRSGYTVHVQ-----CNASK 185

QY 171 SHSCEKVPPTLPQFTMGPNVKINVLEDFEGENPQ-----VPKLEKSISGTSPKREHLP- 223

Db 186 FHQSGSLGVFAIPEFCFAGSDTQRYTSYANANPGEKGGKFFYAQFNKDTAVTSPKREFCPV 245

QY 224 -----LAVGIAF 230

Db 246 DYLLGCGVLIGNAF 259 ; : | | |

Search completed: July 13, 2002, 10:46:37
Job time: 3072 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2002, 09:53:40 ; Search time 24.32 Seconds
(without alignments)
239.033 Million cell updates/sec

Title: US-09-733-756-2
Perfect score: 1301
Sequence: 1 MAAAPLLLLLLLLVLPVLLPL.....REHLPLAVGIAFFLMTFLAS 238

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA: *
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1301	100.0	238	1	US-08-240-124-2
2	1301	100.0	238	1	US-08-453-943-2
3	1301	100.0	238	2	US-09-057-121-2
4	1301	100.0	238	4	US-09-358-734-2
5	1261	96.9	234	1	US-08-299-567-5
6	437	33.6	209	1	US-08-455-001-2
7	437	33.6	209	4	US-08-308-814-2
8	437	33.6	209	5	PCT-US95-11869-2
9	437	33.6	213	1	US-09-609-324A-10
10	437	33.6	213	2	US-08-920-440B-10
11	437	33.6	213	4	US-09-173-492-10
12	437	33.6	213	4	US-09-173-133-10
13	421	32.4	200	1	US-08-455-001-4
14	421	32.4	200	5	PCT-US95-11869-4
15	415.5	31.9	228	1	US-08-442-248-4
16	415.5	31.9	228	1	US-08-440-815-4
17	415.5	31.9	228	3	US-08-379-802-2
18	415.5	31.9	228	3	US-09-048-129-2
19	415.5	31.9	228	4	US-09-048-079-2
20	415.5	31.9	228	4	US-08-486-449-4
21	415.5	31.9	228	5	PCT-US95-15781-5
22	403	31.0	184	1	US-09-609-324A-2
23	403	31.0	184	2	US-08-920-440B-2
24	403	31.0	184	4	US-09-173-492-2
25	403	31.0	184	4	US-09-173-133-2
26	403	31.0	184	4	US-09-165-533-2
27	403	31.0	184	5	PCT-US95-12779-2

28	403	31.0	184	5	PCT-US95-15781-2	Sequence 2, Appli
29	390	30.0	205	1	US-08-321-162-2	Sequence 2, Appli
30	390	30.0	205	1	US-08-448-736-1	Sequence 1, Appli
31	390	30.0	205	1	US-08-441-216-2	Sequence 2, Appli
32	390	30.0	205	1	US-08-452-779-1	Sequence 1, Appli
33	390	30.0	205	1	US-08-299-567-4	Sequence 4, Appli
34	390	30.0	205	2	US-08-445-065-1	Sequence 1, Appli
35	390	30.0	205	2	US-08-445-065-12	Sequence 12, Appli
36	390	30.0	205	3	US-08-959-524-1	Sequence 1, Appli
37	390	30.0	205	3	US-08-959-524-12	Sequence 12, Appli
38	364.5	28.0	179	1	US-08-455-001-5	Sequence 5, Appli
39	364.5	28.0	179	5	PCT-US95-11869-5	Sequence 5, Appli
40	364.5	28.0	201	1	US-08-240-124-4	Sequence 4, Appli
41	364.5	28.0	201	1	US-08-453-943-4	Sequence 4, Appli
42	364.5	28.0	201	2	US-09-057-121-4	Sequence 4, Appli
43	364.5	28.0	201	4	US-09-358-734-4	Sequence 4, Appli
44	334	25.7	104	1	US-09-609-324A-8	Sequence 8, Appli
45	334	25.7	104	2	US-08-920-440B-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-240-124-2
; Sequence 2, Application US/08240124
; Patent No. 5516658
; GENERAL INVENTOR:
; APPLICANT: BECKMANN, M. P.
; TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
; TITLE OF INVENTION: RECEPTOR HEK
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNEX CORPORATION
; STREET: 51 UNIVERSITY STREET
; CITY: SEATTLE
; STATE: WASHINGTON
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,124
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,132
; FILING DATE: 03-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/114,426
; FILING DATE: 30-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/109,745
; FILING DATE: 20-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SEESE, KATHRYN A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2814-C
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-240-124-2


```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-057-121-2

```

Query Match	100.0%;	Score 1301;	DB 2;	Length 238;
Best Local Similarity	100.0%;	Pred. No. 2.7e-128;		
Matches 238;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	MAAAPLLLLLLVVPVPLPLLAQGGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60			
Db 1	MAAAPLLLLLLVVPVPLPLLAQGGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD 60			
QY 61	IYCPHYNSSGVGPGAGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECNRPHAPHSPI 120			
Db 61	IYCPHYNSSGVGPGAGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECNRPHAPHSPI 120			
QY 121	KFSEKFORYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKFVCCASTSHSGEKPVPT 180			
Db 121	KFSEKFORYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKFVCCASTSHSGEKPVPT 180			
QY 181	LPOFTMGPNVKINVLEDFEGENPQVPKLEKISGTSPKREHLPLAVGIAFFLMTFLAS 238			
Db 181	LPOFTMGPNVKINVLEDFEGENPQVPKLEKISGTSPKREHLPLAVGIAFFLMTFLAS 238			

```

RESULT      4
US-09-358-734-2
; Sequence 2, Application US/09358734
; Patent No. 6274117
; GENERAL INFORMATION:
; APPLICANT: BECKMANN, M. P.
; APPLICANT: CERRETTI, DOUGLAS P.
; TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
; TITLE OF INVENTION: RECEPTOR HEK
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNEX CORPORATION
; STREET: 51 UNIVERSITY STREET
; CITY: SEATTLE
; STATE: WASHINGTON
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/358,734
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/240,124
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/114,426
; FILING DATE: 30-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/109,745
; FILING DATE: 20-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SEESE, KATHRYN A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2814-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids

```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-358-734-2

```

	Query Match	100.0%;	Score 1301;	DB 4;	Length 238;
	Best Local Similarity	100.0%;	Pred. No. 2.7e-128;		
	Matches 238;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAAAPLLLLLLVPVPLLLAQQPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD	60		
Db	1	MAAAPLLLLLLVPVPLLLAQQPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLD	60		
QY	61	IYCPHYNSSGVGPGAGPGGGAEQVLYVMVSRNGYRTCNASQGFKRWECNRPHAPHSPI	120		
Db	61	IYCPHYNSSGVGPGAGPGGGAEQVLYVMVSRNGYRTCNASQGFKRWECNRPHAPHSPI	120		
QY	121	KFSEKFQRYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKVFVCCASTSHSGEKPVPT	180		
Db	121	KFSEKFQRYSAFSLGYEFHAGHEYYYISTPTHNLHWKCLRMKVFVCCASTSHSGEKPVPT	180		
QY	181	LPOFTMGPNVKINVLEDFEGENPQVPKLEKSISGTSPKREHLPLAVGIAFFLMTFLAS	238		
Db	181	LPOFTMGPNVKINVLEDFEGENPQVPKLEKSISGTSPKREHLPLAVGIAFFLMTFLAS	238		

```

RESULT      5
US-08-299-567-5
; Sequence 5, Application US/08299567
; Patent No. 5747033
; GENERAL INFORMATION:
; APPLICANT: Davis, et al.
; TITLE OF INVENTION: METHOD OF ENHANCING THE BIOLOGICAL
; TITLE OF INVENTION: ACTIVITY OF EPH FAMILY LIGANDS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10591-6707
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,567
; FILING DATE: 01-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kempler, Gail M.
; REGISTRATION NUMBER: 32,143
; REFERENCE/DOCKET NUMBER: REG 290
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
;
US-08-299-567-5

```

Query Match 96.9%; Score 1261; DB 1; Length 234;
Best Local Similarity 97.9%; Pred. No. 4e-124;
Matches 233; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

QY 1 MAAAPLLLLLLLLVPVPLLLAQQPGGALGNRHAVYWNSSNQHLRRREGYTVQVNVNDYLD 60
Db 1 MAAAPLLLLLLLLVPVPLLLAQQPGGALGNRHAVYWNSSNQHLRRREGYTVQVNVNDYLD 60
QY 61 IYCPHYNSSGVGPGAGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECPHAPHSPI 120
Db 61 IYCPHYNSS----GAGPGPGGGAEQYVLYMVSRNGYRTCNASQGFKRWECPHAPHSPI 116
QY 121 KPSEKFORYSAFSLGYEFHAGHEYIYSTPHNLHWKCLRMKVFCVCCASTSHSGEKPVP 180
Db 117 KPSEKFORYSAFSLGYEFHAGHSYIYSTPHNLHWKCLRMKVFCVCCASTSHSGEKPVP 176
QY 181 LPQFTMGPNVKINVLDFEGENQVPKLEKSISGTSPKREHLPLAVGIAPFLMTFLAS 238
Db 177 LPQFTMGPNVKINVLDFEGENQVPKLEKSISGTSPKREHLPLAVGIAPFLMTFLAS 234
RESULT 6
US-08-455-001-2
; Sequence 2, Application US/08455001
; Patent No. 5795734
; GENERAL INFORMATION:
; APPLICANT: Flanagan, John G.
; APPLICANT: Cheng, Hwai-Jong
; TITLE OF INVENTION: EPH Receptor Ligands, and Uses Related
; TITLE OF INVENTION: Thereto
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,001
; FILING DATE: 31 MAY 1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-011CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-001-2
Query Match 33.6%; Score 437; DB 1; Length 209;
Best Local Similarity 50.5%; Pred. No. 5.4e-38;
Matches 96; Conservative 25; Mismatches 47; Indels 22; Gaps 7;
QY 2 AAAPLLLLLLVPVPLLLA--QGPGGALGNRHAVYWNSSNQHLRRE-----GYTVQV 53
Db 4 AQRPLLLLLL----LLPLRARNEDPARANADRYAVYWNRSNPRFQVSAVGGGGYTV 59
QY 54 NVNDYLDIYCPHYNSSGVGPGAGPGGGAEQYVLYMVSRNGYRTCNASQ-GFKRWE 112
Db 60 SINDYLDIYCPHY-----GAPLPPAERMERYILYMVNGEGHASCDCRQGFKR 112
QY 113 PHAPSPKFKSEKFORYSAFSLGYEFHAGHEYIIS-TPTHNLHWKCLRMKVFCVCCAS-T 170
; Sequence 2, Application US/08455001

Db 113 PAAPGGPLKFSEKFQFTFSLGFEFPRGHEYYIISATPPNLVDRPCLRLKVVYVRPTNET 172
QY 171 SHSGEKPVP 180
Db 173 LYEAPPEIFT 182
RESULT 7
US-08-308-814-2
; Sequence 2, Application US/08308814
; Patent No. 6268476
; GENERAL INFORMATION:
; APPLICANT: Flanagan, John G.
; APPLICANT: Cheng, Hwai-Jong
; TITLE OF INVENTION: EPH Receptor Ligands, and Uses Related
; TITLE OF INVENTION: Thereto
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/308,814
; FILING DATE: 19-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-308-814-2
Query Match 33.6%; Score 437; DB 4; Length 209;
Best Local Similarity 50.5%; Pred. No. 5.4e-38;
Matches 96; Conservative 25; Mismatches 47; Indels 22; Gaps 7;
QY 2 AAAPLLLLLLVPVPLLLA--QGPGGALGNRHAVYWNSSNQHLRRE-----GYTVQV 53
Db 4 AQRPLLLLLL----LLPLRARNEDPARANADRYAVYWNRSNPRFQVSAVGGGGYTV 59
QY 54 NVNDYLDIYCPHYNSSGVGPGAGPGGGAEQYVLYMVSRNGYRTCNASQ-GFKRWE 112
Db 60 SINDYLDIYCPHY-----GAPLPPAERMERYILYMVNGEGHASCDCRQGFKR 112
QY 113 PHAPSPKFKSEKFORYSAFSLGYEFHAGHEYIIS-TPTHNLHWKCLRMKVFCVCCAS-T 170
Db 113 PAAPGGPLKFSEKFQFTFSLGFEFPRGHEYYIISATPPNLVDRPCLRLKVVYVRPTNET 172
QY 171 SHSGEKPVP 180
Db 173 LYEAPPEIFT 182
RESULT 8
PCT-US95-11869-2
; Sequence 2, Application PC/TUS9511869


```

Db      64  SINDYLDIYCPHY-----GAPLPFAERMEHYVLYMVNGEGHASCDDHRRQGRFKRWECNR 116
QY      113  PHAPHSPIKFSEKFORYSAFSLGVFEFHAGHEYIIS-TPTHNLHWKCLRMKVFYCCAS-T 170
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      117  PAAPGGDLKFSEKQLFTPFSLGFEFRPGHEYIISATPPNAVDRCPLRLKLVYVRPTNET 176
QY      171  SHSGEKPVPPT 180
      :  :  |  |  |
Db      177  LYEAEPIPT 186

```

```

RESULT      11
US-09-173-492-10
; Sequence 10, Application US/09173492
; Patent No. 6194172
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; TITLE OF INVENTION: Cytokine Designated LERK-6
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: System 7.6
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/173,492
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/320,440
; FILING DATE: 29-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C.
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2826-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 213 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-173-492-10

```

Query Match	33.6%;	Score 437;	DB 4;	Length 213;
Best Local Similarity	50.5%;	Pred. No. 5.6e-38;		
Matches	96;	Conservative 24;	Mismatches 52;	Indels 18; Gaps 7;

QY	2	AAAPLL-LLLLLVVPVPLLPLL-AQGGGALGNRRHAVYWNSSNQHLRR-----EGYTVQV	53
		: : : :	:
Db	4	AQRPLLPLLLLLLPLPPPPFARAEDAARANSRDYAVYWNRSNPRFHAGAGDGGGYTVEV	63
QY	54	NVNDYLDIYCPHYNSSGVPGAGCPGGGAEQVLYVMVSRNGYRTCNASQ-GFKRWE CNR	112
		: : : : : : :	
Db	64	SINDYLDIYCPHY-----GAPLPPAERMEHYVLYVMVNGEGHASCDDRQRFKRWECNR	116
QY	113	PHAPSPIKFSEKFORYSAFSLGYEFHAGHEYIYIS-TPTHNLHWKCLRMKVFFVCCAS-T	170
		: : : : :	:
Db	117	PAAPGGPLKFSK FQ LFTPFSLGF EFRPGHEYIYISATPPNAVDRPCLRLKVYVRPTNET	176
QY	171	SHSGEKPVPPT	180
		: :	
Db	177	LYEAPEPIPT	186

RESULT 12
US-09-173-133-10
; Sequence 10, Application US/09173133
; Patent No. 6232447
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; TITLE OF INVENTION: Cytokine Designated LERK-6
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: System 7.6
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/173,133
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,440
; FILING DATE: 29-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C.
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2826-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 213 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-173-133-10

```

Query Match          33.6%;   Score 437;   DB 4;   Length 213;
Best Local Similarity 50.5%;   Pred. NO. 5.6e-38;
Matches 96; Conservative 24; Mismatches 52; Indels 18; Gaps 7;

QY  2  AAAPLL-LLLLVPVPLLPLL-AQGGGALGNRHAVYWNSSNQHLRR-----EGYTVQV 53
      | ||| |||||:| | | | | : | : ||:|||| | | ||||:|
Db  4  AQRPLLPLLLLLPLPPPPFARAEDAARANSRDYAVYWNRSNPRFHAGAGDDGGGYTVEV 63
      | ||| ||||| || | | | | | | | | | | | | | | | | | | | | | |
QY  54  NVNDYLDIYCPHYNSSGVGPAGPGPGGAEQVLYVMVSRNGYRTCNASQ-GFKRWEQNR 112
      :||| ||||| || | | | | | | | | | | | | | | | | | | | | | |
Db  64  SINDYLDIYCPHY-----GAPLPPAERMEHVLYMVNGEGHASCDDRQGRFKRWEQNR 116
      | ||| ||||| || | | | | | | | | | | | | | | | | | | | | | |
QY  113 PHAPSPITKSEKFORYSAFSLGYEFHAGHEYIYIS-TPTHNLHWKCLRMKVTVCCAS-T 170
      | || | ||||| || : ||||| || | ||||| || : | | | : || : |
Db  117 PAAPGGPLKSEKQFLTFFSLGFERFPGHEYIYISATPPNAVDRPCLRLKVYVRPTNET 176
      | || | ||||| || : ||||| || | ||||| || : | | | : || : |

QY  171 SHSGEKPVPPT 180
      : | : | |
Db  177 LYEAPPIPT 186

RESULT 13
US-08-455-001-4
; Sequence 4, Application US/08455001
; Patent No. 5795734
; GENERAL INFORMATION:
; APPLICANT: Flanagan, John G.
; APPLICANT: Cheng, Hwai-Jong

```

```

; TITLE OF INVENTION:  EPH Receptor Ligands, and Uses Related
; TITLE OF INVENTION:  Thereto
; NUMBER OF SEQUENCES:  5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  LAHIVE & COCKFIELD
; STREET:  60 State Street
; CITY:  Boston
; STATE:  MA
; COUNTRY:  USA
; ZIP:  02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  PC-DOS/MS-DOS
; SOFTWARE:  ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/08/455,001
; FILING DATE:  31 MAY 1995
; CLASSIFICATION:  800
; ATTORNEY/AGENT INFORMATION:
; NAME:  Vincent, Matthew P.
; REGISTRATION NUMBER:  36,709
; REFERENCE/DOCKET NUMBER:  HMI-011CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  (617) 227-7400
; TELEFAX:  (617) 227-5941
; INFORMATION FOR SEQ ID NO:  4:
; SEQUENCE CHARACTERISTICS:
; LENGTH:  200 amino acids
; TYPE:  amino acid
; TOPOLOGY:  linear
; MOLECULE TYPE:  protein
; US-08-455-001-4

```

Query Match	32.4%;	Score 421;	DB 1;	Length 200;
Best Local Similarity	47.5%;	Pred. NO. 2.4e-36;		
Matches 86;	Conservative 29;	Mismatches 50;	Indels 16;	Gaps 5;

Qy	3	AAPLLLLLLVPVPLPLLAQPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLDIY	62
Db	6	AAALLAAIVGCV-----WSDDPGKVISDRYAVYWNRSNPRFHRGDTYVEVSINDYLDIY	60
Qy	63	CPHYNSSGVGPAGPGGGAEQYVLYMVSRNGYRTCNASQ-GFKRWE CNRPHAPHSPIK	121
Db	61	CPHYEE-----PLAERMERYVLYMVNYEGHASCDDHRQKGFKRWE CNRPDPSGGLK	112
Qy	122	FSEKFORYSAFSLGYEFHAGHEYYYIS-TPTHNLHWKCLRMKMFVCCASTS-HSGEKPVP	179
Db	113	FSEKFQLETFPSLGFEPFGHEYYYISASPPNVVDRPCLKLKVYVRPTNDSLYESPEPIF	172
Qy	180	T 180	
Db	173	T 173	

RESULT 14
PCT-US95-11869-4
; Sequence 4, Application PC/TUS9511869
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: EPH Receptor Ligands, and Uses Related
; TITLE OF INVENTION: Thereto
; NUMBER OF SEQUENCES: 5
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/11869
; FILING DATE: 19-SEP-1995
; ATTORNEY/AGENT INFORMATION:

```

; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-011CPPC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 200 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-11869-4

```

Query Match 32.4%; Score 421; DB 5; Length 200;
Best Local Similarity 47.5%; Pred. NO. 2.4e-36;
Matches 86; Conservative 29; Mismatches 50; Indels 16; Gaps 5;

QY	3	AAPLLLLLLVPVPLP	LAAQGPGGALGNRHAVYWNSSNQHLRRREGYTVOVNVDYLDIY	62
Dd	6	AAALLAAIVGCV----	WSDDPCKVISDRYAVYWNRNPRFHRGDYTVVEVSINDYLDIY	60
QY	63	CPHYNSSGVGPAGPGGGAEQVLYMVSRNGYRTCNASQ-GFKRWECNRPHAPHSPIK	121	
Dd	61	CPHYEE-----PLPAERMERYLYMVMNYEGHASCDDRQKGFGRWECNRPDSPSGPLK	112	
QY	122	FSEKFQRYSAFSLGVEFHAGHEYYYYIS-TPTHNLHWKCLRMKVFCCASTS-HSGEKPVP	179	
Dd	113	FSEKFQLFTPFSLGFEFRPGHEYYYYISASPNNVDRPCCLKKVYVRPTNDLSYESPEIF	172	
QY	180	T	180	
Dd	173	T	173	

RESULT 15
US-08-442-248-4
; Sequence 4, Application US/08442248
; Patent No. 5759863
; GENERAL INFORMATION:
; APPLICANT: Caras, Ingrid W.
; APPLICANT: Winslow, John W.
; TITLE OF INVENTION: AL-1 Neurotrophic Factor
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/442,248
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/330128
; FILING DATE: 27-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: 920C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8674
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ IDNO: 4:

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2002, 09:49:45 ; Search time 57.73 Seconds
(without alignments)
457.917 Million cell updates/sec

Title: US-09-733-756-2
Perfect score: 1301
Sequence: 1 MAAAPLLLLLLLLVPVLLPL.....REHPLAVGIAFFLMTFLAS 238

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1301	100.0	238	16 AAR71481	Human hek-L protel
2	1266	97.3	234	16 AAR82605	Eph transmembrane
3	437	33.6	209	17 AAR94766	Mouse Eph receptor
4	437	33.6	209	19 AAW71006	Amino acid sequenc
5	437	33.6	213	20 AAY06822	Human LERK-6 polyp
6	426.5	32.8	335	22 ABG27837	Novel human diagno
7	421	32.4	200	17 AAR94767	Chicken Eph recept
8	421	32.4	200	19 AAW71007	Amino acid sequenc
9	415.5	31.9	228	17 AAR97854	Human AL-1, a liga
10	415.5	31.9	228	17 AAW02586	Lerk-7 protein. H
11	415.5	31.9	228	17 AAW00035	HEK4 binding prote

12	403	31.0	184	17 AAW02587	Lerk-6 protein. M
13	403	31.0	184	17 AAR91283	LERK-6 protein. M
14	403	31.0	184	20 AAY06820	Murine LERK-6 poly
15	390	30.0	205	13 AAR23895	B61 protein sequen
16	390	30.0	205	15 AAR53634	EBP. Homo sapiens
17	390	30.0	205	16 AAR82604	Eph transmembrane
18	390	30.0	205	18 AAW04632	eck receptor bindi
19	386	29.7	205	22 AAB50990	Human PRO202 prote
20	386	29.7	218	21 AAB54058	Human pancreatic c
21	384	29.5	180	18 AAW18082	Truncated eck rece
22	382	29.4	171	18 AAW18081	Truncated eck rece
23	380	29.2	167	18 AAW18080	Truncated eck rece
24	379.5	29.2	160	18 AAW18079	Truncated eck rece
25	377.5	29.0	151	18 AAW18078	Truncated eck rece
26	377.5	29.0	204	22 AAU12315	Human PRO202 polyp
27	377.5	29.0	204	22 AAB50974	Human PRO202 prote
28	364.5	28.0	179	17 AAR94768	EPH receptor ligan
29	364.5	28.0	179	19 AAW71008	Generic sequence f
30	364.5	28.0	201	16 AAR71482	Human hek-L protei
31	334	25.7	104	17 AAR91284	LERK-6 exon polype
32	334	25.7	104	20 AAY06821	Amino acid sequenc
33	295	22.7	58	21 AAG03418	Human secreted pro
34	180	13.8	340	18 AAW17081	EPH family ligand
35	180	13.8	340	19 AAW46615	Human transmembran
36	179	13.8	340	18 AAW33699	AL-2-short (AL-2s)
37	179	13.8	340	18 AAW31544	Human cytokine Ler
38	179	13.8	340	18 AAW10637	NLERK2 ligand for
39	179	13.8	455	18 AAW33698	AL-2-long (AL-2l)
40	177	13.6	658	21 AAY96782	Ephrin-B2-Ephrin-B
41	171.5	13.2	308	17 AAW06334	Ligand #2 for rece
42	171.5	13.2	308	17 AAR94656	Ligand for recepto
43	171.5	13.2	333	17 AAW06337	Full length ligand
44	171.5	13.2	333	17 AAR94655	Ligand for recepto
45	171.5	13.2	333	17 AAR92743	Human hepatoma tra

ALIGNMENTS

RESULT 1
AAR71481
ID AAR71481 standard; Protein; 238 AA.
XX
AC AAR71481;
XX
DT 03-OCT-1995 (first entry)
XX
DE Human hek-L protein.
XX
KW Ligand; cell surface; tyrosine kinase receptor; tumorigenesis;
KW immunogen.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT Protein /note= "signal peptide"
XX 20..238
XX WO9506065-A.
XX 02-MAR-1995.

XX 17-AUG-1994; 94WO-US09282.
XX
PR 20-AUG-1993; 93US-0109745.
PR 30-AUG-1993; 93US-0114426.
PR 03-DEC-1993; 93US-0161132.
PR 09-MAY-1994; 94US-0240124.
XX
PA (IMMV) IMMUNEX CORP.
XX Beckmann MP, Cerretti DP;
PI

CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 335 AA;

Query Match 32.8%; Score 426.5; DB 22; Length 335;
Best Local Similarity 39.2%; Pred. NO. 7.8e-35;
Matches 100; Conservative 39; Mismatches 69; Indels 47; Gaps 11;

QY 1 MAAAPLLLLLLVPVPLLLAQQGG-ALGNRHAVYWNSSNQHLRREGYTVQVNVNDYL 59

Db 106 lgvapvemltlvflvmcvfsqdpqskavadyavywnssnprfgrgdyhidvcindy1 165

QY 60 DIYCPHYNSSGVGPGAGPGGGAEQYVLYMVSRYGRTTC-NASQGFKRWECNRPHAPHS 118

Db 166 dvfcphyeds-----vpedkteryvlymvmnfdgysacdhtskgfkrcwecnrphspng 217

QY 119 PIKFSEKFORYSAFSLGYEFHAGHEYYIISTP-THNLHWKCLRMKVFEV----CCASTSHS 173

Db 218 plkfsekqlftpslglfepfprgryfyissalpdngrrscldklkvfrptnscmktigv 277

QY 174 GEKVPVPTLPQFTMGPNVKINLEDFE-----GEN-PQVPKLEKSISGTSPKREH 221

Db 278 hdr-----vfdvndkve-nslepaddtvhesaeprsgenaagtpr-----psr-- 320

QY 222 LPLAVGIAFFLMTFL 236

Db 321 ---llailllaml 332

RESULT 7

ID AAR94767 standard; Protein; 200 AA.

XX AAR94767;

DT 02-JUL-1996 (first entry)

XX Chicken EPH receptor ligand Elf-1.

DE Elf-1; EPH receptor ligand; dementia; tachycardia; therapy;

XX diagnosis; transgenic animal.

OS Gallus sp.

XX Key Location/Qualifiers

FT Peptide 1..21

FT /label= Sig_peptide

FT 22..200

FT /label= Mat_protein

FT 61..150

FT /label= Cys4_motif

FT 33..157

FT /label= Core_sequence_motif

XX WO9609384-A1.

PN 28-MAR-1996.

PD 19-SEP-1995; 95WO-US11869.

XX 27-FEB-1995; 95US-0393462.

PR 19-SEP-1994; 94US-0308814.

XX (HARD) HARVARD COLLEGE.

XX WPI; 1996-188446/19.

DR N-PSDB; AAT15009.

XX Murine and chicken EPH receptor ligand, Elf-1 - useful in diagnosis

PT and treatment of disorders associated with the Elf-1 gene, e.g.

PT dementia, tachycardia , etc.

XX Claim 1; Page 88-89; 107pp; English.

XX A novel chicken EPH receptor ligand, Elf-1 (AAR94767), is involved in
CC the formation and maintenance of ordered spatial arrangements of
CC differentiated tissue. It is the product of a cDNA clone (AAT15009)
CC obtd. from an embryo cDNA expression library. This cDNA is used for
CC the prodn. of recombinant Elf-1, which can be used to modulate
CC proliferation, survival and/or differentiation of cells and tissues,
CC and to stimulate or antagonise intracellular signal transduction
CC pathways mediated by the EPH-type receptor.

XX Sequence 200 AA;

Query Match 32.4%; Score 421; DB 17; Length 200;

Best Local Similarity 47.5%; Pred. NO. 1.4e-34;

Matches 86; Conservative 29; Mismatches 50; Indels 16; Gaps 5;

QY 3 AAPLLLLLLVPVPLLLAQQGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLDIY 62

Db 6 aaallaaivgcv-----wsddpgkvisdryavywnrsnprfhrgrdytvevsindyldiy 60

QY 63 CPHYNSSGVGPGAGPGGGAEQYVLYMVSRYGRTCNASQ-GFKRWECNRPHAPHSPIK 121

Db 61 cphyee-----plpaermeryvlymvmnyeghascdhrqkgfkrcwecnrpsgplk 112

QY 122 FSEKFORYSAFSLGYEFHAGHEYYIIS-TPTHNLHWKCLRMKVFCVCCASTS-HSGEKVPV 179

Db 113 fsekqlftpslglfepfprgryfyissalpdngrrscldklkvfrptnscmktigv 172

QY 180 T 180

Db 173 t 173

RESULT 8

ID AAW71007 standard; Protein; 200 AA.

XX AAW71007;

DT 20-OCT-1998 (first entry)

XX Amino acid sequence of an avian Elf-1 protein.

DE Chicken; EPH receptor ligand; Elf-1; mek-4; sek-AP;
XX tyrosine kinase ligand; B61; LERK-2; proliferation; differentiation;
KW intracellular signalling; increased; survival; neuronal cell;
KW neuron survival; treatment; Alzheimer's; Parkinson's; lymphatic tumour;
KW artificial liver; cartilage; bone formation.

OS Gallus sp.

XX Key Location/Qualifiers

FT Misc-difference 61..150

FT /note= "contains a Cys4 motif"

FT Misc-difference 35..157

FT /note= "contains a core sequence motif"

XX US5795734-A.

PN 18-AUG-1998.

XX 31-MAY-1995; 95US-0455001.

XX

PR 31-MAY-1995; 95US-0455001.
PR 19-SEP-1994; 94US-0308814.
PR 27-FEB-1995; 95US-0393462.

XX (HARD) HARVARD COLLEGE.

XX Cheng H, Flanagan JG;

PI WPI; 1998-466665/40.
XX N-PSDB; AAV42927.

PT Nucleic acid encoding Elf-1 protein that binds to EPH-type receptor
PT - for production of Elf-1 protein, useful for regulating
PT proliferation, differentiation, and survival of cells

XX Claim 1; Columns 75-78; 53pp; English.

XX The present sequence represents an avian EPH receptor ligand designated
CC Elf-1. This ligand can bind to both mek-4 and sek-AP. Elf-1 is a
CC tyrosine kinase ligand, which is linked to the membrane through a
CC phosphatidylinositol linkage. It shares some homology to 2 other EPH
CC receptor ligands, B61 and LERK-2. The Elf-1 protein modulates
CC proliferation, differentiation and survival of EPH receptor-expressing
CC cells by stimulating or antagonising intracellular signalling mediated
CC by the EPH receptor. Typical of many potential applications are
CC increasing survival of neuronal cells in culture (e.g. where intended for
CC transplantation), also therapeutically in increase neuron survival
CC (e.g. treatment of Alzheimer's or Parkinson's diseases), to prevent
CC nervous system and lymphatic tumours, to induce differentiation of
CC hepatocytes to form an artificial liver, to induce cartilage and bone
CC formation.

XX Sequence 200 AA;

Query Match 32.4%; Score 421; DB 19; Length 200;
Best Local Similarity 47.5%; Pred. No. 1.4e-34;
Matches 86; Conservative 29; Mismatches 50; Indels 16; Gaps 5;

QY 3 AAPLLLLLLVPVPLLLLAQPGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLDIY 62

Db 6 aaallaavgvcv-----wsddpgkvisdryavywnrsnprfhrqdytvevsindyldiy 60

QY 63 CPHYNSSGVGPAGPGPGGAEQYVLYMVSRNGYRTCNASQ-GFKRWECNRPHAPHSPIK 121

Db 61 cphyee-----plpaermeryvlymnyeghascdhrqgfkwrwecnrpsgplk 112

QY 122 FSEKFORYSAPSLGYEFHAGHEYVYIS-TPTNHLHWKCLRMKVFVCCASTS-HSGEKPVP 179

Db 113 fsekflftpslglfepfpgheyysasppnvvdrcplklkyvrptndslsyespepif 172

QY 180 T 180

Db 173 t 173

RESULT 9

AAAR97854

ID AAR97854 standard; Protein; 228 AA.

XX AAR97854;

DT 05-JAN-1997 (first entry)

XX Human AL-1, a ligand for eph-related tyrosine kinase receptor REK7.

DE AL-1; REK7; eph-related tyrosine kinase receptor; ligand;

XX neurotrophic factor; neuropathy; angiogenesis; therapy; diagnosis.

OS Homo sapiens.

XX key Location/Qualifiers

FH Peptide 1..20

FT

FT Modified-site /label= Sig_peptide
FT 37..39
FT /label= Glycosylation
FT /note= "potential N-glycosylation site"
FT 162..164
FT /label= Glycosylation
FT /note= "potential N-glycosylation site"
FT 203..204
FT /label= GPI
FT /note= "potential attachment site for
FT glycoposphatidyl-inositol"
FT 214..228
FT /label= Hydrophobic_domain

XX WO9613518-A1.

XX 09-MAY-1996.

XX 26-OCT-1995; 95WO-US14016.

XX 07-JUN-1995; 95US-0486449.

XX 27-OCT-1994; 94US-0330128.

XX (GETH) GENENTECH INC.

XX Caras IW, Winslow JW;

XX WPI; 1996-239448/24.

XX N-PSDB; AAT18897.

XX AL-1 neurotrophic factor, an eph-tyrosine kinase ligand - used in

XX treatment and diagnosis of neuronal disorders and

XX angiogenesis-related conditions.

XX Claim 1; Page 54; 75pp; English.

XX Human AL-1 (AAR97854) neurotrophic factor is a ligand for eph-related
CC tyrosine kinase receptor REK7 (AAR97853). Its amino acid sequence
CC was deduced from a cDNA clone (AAT18897) isolated from a human foetal
CC brain cDNA library. Recombinant, mature AL-1 can be produced in
CC transformed host cells. It may be useful in promoting the
CC development, maintenance or regeneration of neurons in vivo, and
CC can be utilised in methods for the diagnosis and/or treatment of
CC neurological disorders. It can also be used to promote or enhance
CC angiogenesis. Antagonists and agonists of AL-1, as well as
CC antibodies raised against AL-1, also have therapeutic applns.

XX Sequence 228 AA;

Query Match 31.9%; Score 415.5; DB 17; Length 228;
Best Local Similarity 39.5%; Pred. No. 6.1e-34;
Matches 98; Conservative 37; Mismatches 66; Indels 47; Gaps 11;

QY 8 LLLLLVPVPLLLLAQPGG-ALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLDIYCPHY 66

Db 6 mltlvflvlwmcvfsqdpqskavadyavywnssnprfgrgdyhidvcindyldvfcphy 65

QY 67 NSSGVGPAGPGPGGAEQYVLYMVSRNGYRTC-NASQGFKRWECNRPHAPHSPIKFESEK 125

Db 66 eds-----vpedkteryvlymvmvfdgysacdhtskgfkwrwecnrphspngplkfsek 117

QY 126 FQRYSAFSLGYEFHAGHEYVYISTP-TNHLHWKCLRMKVVFV-----CCASTSHSGEKPVP 180

Db 118 fqlftpslglfepfpgreyfysaipdngrsrclklkvfvrptnscmktlgyvhdr----- 173

QY 181 LPQFTMGPNVKINVLEDFE-----GEN-PQVPKLEKSISGTSPPKREHLPLAVGI 228

Db 174 --vfdvndkve-nslepaddtvhesaepsgenaagtpri-----psr-----llai 217

QY 229 AFFLMTFL 236

Db 218 llfllaml 225

RESULT	10
AAW02586	
ID	AAW02586 standard; Protein; 228 AA.
XX	
AC	AAW02586;
XX	
DT	28-NOV-1996 (first entry)
XX	
DE	Lerk-7 protein.
XX	
KW	Lerk-6; hek; elk; cell surface receptor; culture; reagent;
KW	neuron; disorder; injury; delivery agent; diagnostic; therapeutic;
KW	Lerk-7; probe; cytokine.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	1..20
FT	/label= signal_peptide
FT	1..228
FT	/label= precursor_protein
FT	1..133
FT	/label= extracellular_receptor-binding_domain
FT	134..183
FT	/label= spacer_region
FT	194..208
FT	/note= "C-terminal stretch of hydrophobic residues"
FT	183
FT	/label= GPI_attachment_site
XX	
PN	WO9617925-A1.
XX	
PD	13-JUN-1996.
XX	
PF	05-DEC-1995; 95WO-US15781.
XX	
PR	01-MAR-1995; 95US-0396946.
PR	06-DEC-1994; 94US-0351025.
XX	
PA	(IMMV) IMMUNEX CORP.
XX	
PI	Cerretti DP;
XX	
DR	WPI; 1996-287171/29.
DR	N-PSDB; AAT32699.
XX	
PT	New isolated human Lerk-7 cytokine - which binds to cell surface
PT	receptors elk, hek and eck, useful for delivering agents to cells or
PT	for treating neural disorders
XX	
PS	Claim 1; Page 37-38; 49pp; English.
XX	
CC	The present sequence is that of human Lerk-7 protein. Lerk-7 is
CC	predicted to be anchored to the cell surface via
CC	glycosyl-phosphatidylinositol (GPI) linkage. A GPI anchor attaches to the
CC	exposed C-terminal amino acid of the processed mature protein, usually
CC	after cleavage upstream, often about 10-12 amino acids, of the N-terminus
CC	of the hydrophobic domain. Lerk-7 coding sequence (AAT32699) was isolated
CC	using a probe derived from the murine Lerk-6 DNA (AAT32700). The Lerk-7
CC	gene and protein can be used for studying the role of Lerk-7 in
CC	conjunction with elk, hek and eck receptors. They can also be used for
CC	delivering diagnostic or therapeutic agents to cells, e.g. cancer cells.
CC	The Lerk-7 proteins can also exhibit neuroprotective or neurotrophic
CC	properties and can be used to treat neural tissue disorders.
XX	
SQ	Sequence 228 AA;
Query Match	
Best Local Similarity	31.9%; Score 415.5; DB 17; Length 228;
Matches	98; Conservative 37; Mismatches 66; Indels 47; Gaps 11;

CC damaged or depleted tissues. Antagonists are useful for cancer
CC treatment. HEK4 BP can also be used to raise antibodies.
XX
SQ Sequence 228 AA;

Query Match 31.9%; Score 415.5; DB 17; Length 228;
Best Local Similarity 39.5%; Pred. No. 6.1e-34;
Matches 98; Conservative 37; Mismatches 66; Indels 47; Gaps 11;

QY 8 LLLLLVPVPLLLAQPGG-ALGNRHAVYWNSSNOHLRREGYTVQVNVNDYLDIYCPHY 66
:| | : : :| | | : :| | | | | :| | :| | | :| | |
Db 6 mltlvflvwmcvfsqdpqskavadyavywnssnprfgrgdyhidvcindvldvfcphy 65

QY 67 NSSGVPAGPGPGGGAEQVLYMVSNGYRTC-NASQGFKRWEKNRPHAPHSPIKFSEK 125
| :| | | | :| | :| | | | :| | :| | | | :| | | |
Db 66 eds-----vpedkterylvymvndfgyssadhtskgfrwecnrphspngplkfsek 117

QY 126 FQYSAFSLGYEFHAGHEYIYSTP-THNLHWKCLRMKVYFV-----CCASTSHSGEKPVP 180
|| :| | | | | :| | | | :| | | | | :| | | | :| | :| |
Db 118 fqifpfslgfeyfrgryfysaipdngrrrsclklkvfvrptnscmktigvhdr---- 173

QY 181 LPQFTMGPNVKINVLDFE-----GEN-PQVPKLEKSISGTSRKREHPLAVGI 228
| :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 174 --vfdvndkve-nslepaddtvhesaepsrgenaagtprl-----psr-----llai 217

QY 229 AFFLMTFL 236
|| :| |
Db 218 lllfllaml 225

RESULT 12
AAW02587
ID AAW02587 standard; Protein; 184 AA.
XX
AC AAW02587;
XX
DT 28-NOV-1996 (first entry)
XX
DE Lerk-6 protein.
XX
KW Lerk-6; hek; elk; cell surface receptor; culture; reagent;
KW neuron; disorder; injury; delivery agent; diagnostic; therapeutic;
KW Lerk-7; probe; cytokine.
XX
OS Mus musculus.
XX
PN WO9617925-A1.
XX
PD 13-JUN-1996.
XX
PF 05-DEC-1995; 95WO-US15781.
XX
PR 01-MAR-1995; 95US-0396946.
PR 06-DEC-1994; 94US-0351025.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Cerretti DP;
XX
DR WPI; 1996-287171/29.
DR N-PSDB; AAT32700.
XX
PT New isolated human Lerk-7 cytokine - which binds to cell surface
PT receptors elk, hek and eck, useful for delivering agents to cells or
PT for treating neural disorders
XX
PS Example 1; Page 35; 49pp; English.
XX
CC The present sequence is that of a murine cytokine, Lerk-6, encoded by
CC AAT32700. A fragment of murine Lerk-6 DNA was isolated by PCR for use as
CC a probe for a human Lerk-6 homologue. The probe however led to the
CC identification of a human Lerk-7 gene (see AAT32699). The Lerk-7 gene

CC and protein can be used for studying the role of Lerk-7 in conjunction
CC with elk, hek and eck receptors. They can also be used for delivering
CC diagnostic or therapeutic agents to cells, e.g. cancer cells. The
CC Lerk-7 proteins can also exhibit neuroprotective or neurotrophic
CC properties and can be used to treat neural tissue disorders.
XX
SQ Sequence 184 AA;

Query Match 31.0%; Score 403; DB 17; Length 184;
Best Local Similarity 50.6%; Pred. No. 8.5e-33;
Matches 82; Conservative 24; Mismatches 40; Indels 16; Gaps 5;

QY 28 ALGNRHAVYWNSSNOHLRRE-----GYTVQVNVNDYLDIYCPHYNSSGVGPAGPGGG 81
| :| | | | | :| | :| | | | :| | :| | | | :| | :| |
Db 3 anadryavywnsrnprfqvsavgdgggytvevsindyldiycphy-----gaplppae 55

QY 82 GAEQYVLYMVSNGYRTCNASQ-GFKRWEKNRPHAPHSPIKFSEKFORYSFSLGYEFHA 140
| :| | | | :| | :| | | | | :| | | | | :| | | :| | :| |
Db 56 rmeryilymvnggeghascdhrqrgfkrwecnrpaapggplkfsekfglftpfslgfeirp 115

QY 141 GHEYYYIS-TPTHNLHWKCLRMKVFCVCCAS-TSHSGEKPVP 180
|| | | | | :| | :| | | | :| | :| | :| | :| | :| |
Db 116 gheyyyisatppnlvdprclrlkvvyrtnetlyeapepift 157

RESULT 13
AAR91283
ID AAR91283 standard; Protein; 184 AA.
XX
AC AAR91283;
XX
DT 10-OCT-1996 (first entry)
XX
DE LERK-6 protein.
XX
KW LERK-6; hek; elk; cell surface receptors; culture; reagent;
KW neurons; disorder; injury; delivery agent; diagnostic; therapeutic.
XX
OS Mus musculus.
XX
PN WO9610911-A1.
XX
PD 18-APR-1996.
XX
PF 04-OCT-1995; 95WO-US12779.
XX
PR 03-OCT-1995; 95US-0538709.
PR 05-OCT-1994; 94US-0318393.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Cerretti DP;
XX
DR WPI; 1996-209575/21.
DR N-PSDB; AAT14009.
XX
PT Isolated DNA encoding cytokine designated LERK-6 which binds to hek
PT and elk cell surface receptors - useful for drug delivery and
PT screening procedures.
XX
PS Claim 6; Page 34; 44pp; English.
XX
CC The LERK-6 polypeptide encoded can be used to isolate cells
CC expressing hek/elk cell surface receptors, or to measure the
CC biological activity of such receptors. The protein may also be used
CC as a delivery agent, taking diagnostic and therapeutic agents to
CC cells expressing such receptors. LERK-6 can also be used as a tissue
CC culture reagent to enhance the viability or prolong the lifespan of
CC the neurons. Neural tissue disorders and injuries may be treated by
CC contact with the polypeptide.
XX
SQ Sequence 184 AA;

Query Match	31.0%;	Score 403;	DB 17;	Length 184;
Best Local Similarity	50.6%;	Pred. No. 8.5e-33;		
Matches 82;	Conservative 24;	Mismatches 40;	Indels 16;	Gaps 5;

QY	28	ALGNRRHAVYWNSSNOHLRRE-----GYTVQVNVNDYLDIYCPHYNSSGVGPAGPGPGG	81
Db	3	anadryavywnrsnprfqvsavdggytvevsindyldiycphy-----gaplppae	55
QY	82	GAEQYVLYMVSRNGYRTCNASQ-GFKRWECCNRPHAPHSPIKFSEKFQRYSAFSLGYEFHA	140
Db	56	rmaryilylmvngedhascdhrqgfkrcwecnrpaapggplkfsekqlftfpfslgfeifr	115
QY	141	GHEYIIIS-TPTHNLHWKCLRMKVFVCCAS-TSHSGEKPVPPT	180
Db	116	gheyyiyisatppnlvdrcplrklkvyvrptnetlyeapepift	157

RESULT 14
 AAY06820
 ID AAY06820 standard; Protein; 184 AA.

AC	AAY06820;
XX	
DT	24-JUN-1999 (first entry)
XX	
DE	Murine LERK-6 polypeptide.
XX	
KW	LERK-6 polypeptide; hek receptor; elk receptor; human; murine; injury;
KW	cell proliferation; neural growth; neural tissue; neurological disease;
KW	neurodegenerative; excitotoxicity.

MUS sp.

WO9910495-A1.

04-MAR-1999.

27-AUG-1998; 98WO-US17772.

29-AUG-1997; 97US-0920440.

(IMMV) IMMUNEX CORP.

Cerretti DP;

WPI; 1999-243567/20.

N-PSDB; AAX32761.

New cytokine designated LERK-6

Claim 6; Page 37-38; 46pp; English.

The invention relates murine and human LERK-6 polypeptides that bind to hek/elk receptors. Host cells transfected or transformed with vectors comprising the LERK-6 nucleic acid sequences are used for the recombinant production of the proteins. LERK-6 polypeptides may be useful in the enhancement, stimulation, proliferation or growth of cells expressing the hek or elk receptor. The ligand and receptor complex may be involved in neural growth, development and/or maintenance. LERK-6 can be used for treating disorders of neural tissue such as injury or neurological diseases, either chronic or acute. LERK-6 may be employed in treating neurodegenerative conditions where there is neural death, excitotoxicity. In addition, they may be administered to a mammal to exert a trophic effect on neural tissue. They can also be used as reagents for those conducting quality assurance studies e.g. to monitor shelf life and stability of elk protein under different conditions. The polypeptides can also be used as carriers for delivering agents attached to cells bearing the elk or hek cell surface receptor. The present sequence represents a murine LERK-6 polypeptide.

Sequence 184 AA;

50

Query Match	31.0%;	Score 403;	DB 20;	Length 184;
Best Local Similarity	50.6%;	Pred. No. 8.5e-33;		
Matches 82;	Conservative 24;	Mismatches 40;	Indels 16;	Gaps 5;

QY	28	ALGNRRHAVYWNSSNQHLRRE-----GYTVQVNVNDYDIYCPHYNSSGVGPGAGPGPGG	81
Db	3	anadryavywnrsnprfqvsavdggggytvevsindydiycphy-----gapippae	55
QY	82	GAEQYVLYMVSRNGYRTCNASQ-GFKRWECNRPHAPHSPIKFSEKQRYSAFSLGYEFHA	140
Db	56	rmeryilylmvnggeghascdhrqrgfkrwecnrpaapgpgplkfsekqlfftpflgfeirp	115
QY	141	GHEYYYIIS-TPTHNLHWKCLRMKVFVCCAS-TSHSGEKKPVPT	180
Db	116	gheyyvisatppnlvdrcplrklkvyrprtnetlyeaepift	157

RESULT	15
--------	----

AAR23895
ID AAR23895 standard; Protein; 205 AA.

AC AAR23895;

DT 06-NOV-1992 (first entry)

DE B61 protein sequence.

Early inflammatory response; marker; antibody; therapy; induction; lipopolysaccharides; cytokines; Interleukin-2; IL-2; TNF.

OS Homo sapiens.

Key	Location/Qualifiers
Peptide	1..18

	FT
	/note= "signal peptide"

PN WO9207094-A.

PD 30-APR-1992.

PF 15-OCT-1991; 91WO-US07704.

PR 16-OCT-1990; 90US-0607741.

PA (UNMI) UNIV MICHIGAN.

PI Dixit VM;

WPI: 1992-167172/20

WEI, 1992 10/17/2/
N-PSDB; AAO24595.

Gene prod. used as marker to detect inflammatory response - by
 aPT detecting the B61 gene in biological fluids or by using
 aPT hybridisation probes

Claim 7; Flg 3; 39pp; English.

The B1 protein sequence was deduced from the cDNA sequence obtained by screening a cDNA library from human umbilical vein endothelial cells with radiolabelled cDNA fragments derived from the 5' end of B61 DNA. The B1 protein has 205 residues (24 kD) comprising a signal sequence and hydrophobic N- and C- terminal regions. The B61 gene is involved in early inflammatory response and serves as a marker. It may be detected by probes or by antibody-based immunoassay of biological fluids such as plasma, CSF or urine. These assays make it possible to predict a worsening in a disease process and allow the quantitative assessment of the magnitude of the inflammatory response. This information will allow the earlier administration of appropriate therapy, thereby shortening the disease process and limiting the patient's exposure to anti-inflammatory/immunosuppressive therapy. B61 induction is rapid and profound, hence

CC it is easily detectable. The B61 response is highly specific to
CC proinflammatory stimuli, being only made by cells exposed to lipo-
CC polysaccharides or cytokines such as IL-2 and TNF and not growth
CC factors or interferon.
XX

SQ Sequence 205 AA;

Query Match 30.0%; Score 390; DB 13; Length 205;
Best Local Similarity 41.8%; Pred. No. 2.1e-31;
Matches 87; Conservative 21; Mismatches 76; Indels 24; Gaps 6;

QY 19 PLLAQGGGALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLDIYCPHYNSSGVGPAGPG 78
Db ||| | :||| ||| | ||| :||| ||| | |

QY 79 PGGAEQVLYVMVSRNGYRTCN-ASQGFKRWEKNRPHAPHSPIKFSEKFQRYSAFSLGYE 137
Db ||| :||| | :| :||| ||| | | ||| ||| :|

QY 138 FHAGHEYVYISTPTNHLHWKCLRMKVVCCEASTSHSGEKPVTLPQFTMGPNVKINVLED 197
Db || || |||| | :||| ||| :|| ||| :|

QY 198 FEGENPQVPKLEKISGTSPKREHLPLA 225
Db |:| | :| :||

Search completed: July 13, 2002, 10:44:47
Job time: 3302 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 13, 2002, 08:40:20 ; Search time 1578.36 Seconds
(without alignments)
14904.833 Million cell updates/sec

Title: US-09-733-756-1
Perfect score: 1743
Sequence: 1 gctgctgctgctgctgc.....aaaaaaaaaaaaaaaaaaaa 1743

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	873.8	50.1	1008	9 AL565095	AL565095 AL565095
2	872.6	50.1	1030	9 AL533153	AL533153 AL533153
C 3	760.8	43.6	798	9 AL563020	AL563020 AL563020
4	718.8	41.2	1013	10 BF969747	BF969747 602272160
5	687.8	39.5	702	10 BG743158	BG743158 602634316
C 6	625.4	35.9	750	9 AL572018	AL572018 AL572018
7	619.6	35.5	1085	11 AK020438	AK020438 Mus muscu
8	615.6	35.3	684	9 AL527972	AL527972 AL527972
9	614.4	35.2	826	10 BF968137	BF968137 602269124
10	578.6	33.2	804	10 BF983120	BF983120 602305829
11	566	32.5	762	10 BG698896	BG698896 602703313
12	553.2	31.7	902	9 AL546822	AL546822 AL546822
13	530.2	30.4	547	10 BG499730	BG499730 602545144
C 14	439.2	25.2	496	10 BF953157	BF953157 CM3-NN118
C 15	416.8	23.9	420	9 BE138904	BE138904 xw97e11.x
C 16	413.4	23.7	435	9 AW189015	AW189015 xk98d10.x
17	409	23.5	535	9 AW786993	AW786993 120664 MA

C 18	401.2	23.0	406	9 AW780303	AW780303 hol10h09.x
19	388.8	22.3	516	10 BE663678	BE663678 147613 MA
20	375.4	21.5	553	10 BE663141	BE663141 126832 MA
C 21	351.4	20.2	353	9 AI272654	AI272654 ql59b04.x
22	350.4	20.1	884	10 BG705182	BG705182 602688106
23	347.8	20.0	848	10 BF542148	BF542148 602068243
24	330.6	19.0	879	10 BF977204	BF977204 602146759
25	320.4	18.4	538	9 AW669878	AW669878 113619 MA
26	289.6	16.6	552	9 AL628407	AL628407 AL628407
27	279.2	16.0	654	10 BI731050	BI731050 603351667
28	276.2	15.8	298	9 BE048061	BE048061 tz47d12.y
C 29	274	15.7	390	9 AW123794	AW123794 UI-M-BH2.
30	269	15.4	523	10 BE648416	BE648416 UI-M-BH2.
31	257.8	14.8	460	9 AW200648	AW200648 da24c06.y
32	247	14.2	481	9 AA388313	AA388313 vc24h03.r
33	246.2	14.1	288	9 AA338749	AA338749 EST43752
34	245.8	14.1	788	10 BI753156	BI753156 603026017
C 35	230.8	13.2	275	10 T06173	T06173 EST04062 Fe
36	227.8	13.1	408	10 BF555727	BF555727 UI-R-A1-d
C 37	222.2	12.7	354	10 BG376619	BG376619 UI-R-CU0-
38	221.2	12.7	332	9 AI508818	AI508818 vc24h03.y
39	220.2	12.6	258	10 BF950361	BF950361 CM3-NN118
40	219.8	12.6	370	10 BE721090	BE721090 188082 MA
41	218.4	12.5	906	10 BF240606	BF240606 601875729
C 42	217.4	12.5	750	9 AW199923	AW199923 da24c06.x
C 43	216.8	12.4	252	9 AI863365	AI863365 tz47d12.x
C 44	215.6	12.4	364	9 AA901053	AA901053 UI-R-A1-d
C 45	206	11.8	324	10 BM386597	BM386597 UI-R-CN1-

ALIGNMENTS

RESULT 1
AL565095/c 1008 bp mRNA linear EST 16-FEB-2001
LOCUS AL565095 LTI_FL015_Brn1 Homo sapiens cDNA clone CS0DN003YL19 3
DEFINITION prime, mRNA sequence.
ACCESSION AL565095
VERSION AL565095.1 GI:12916129
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1008)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES source

Location/Qualifiers
1..1008
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DN003YL19"
/clone_lib="LTI_FL015_Brn1"
/sex="male"
/tissue_type="Adult brain"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies.
Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 209 a 262 c 325 g 195 t 17 others
ORIGIN

[illegible]

```

RESULT      3
AL563020/c
LOCUS
DEFINITION
AL563020 LTI_NFL003_NBC3 Homo sapiens CDNA clone CS0DC027YP19 3
prime, mRNA sequence.
AL563020
AL563020.1 GI:12912020
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 798)
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..798
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC027YP19"
/clone_lib="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source

```

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>"

BASE COUNT	ORIGIN
164 a	203 c
257 g	165 t
9 others	

Query Match	43.6%	Score 760.8;	DB 9;	Length 798;
Best Local Similarity	97.9%	Pred. No. 3.7e-94;		

QY	745	tggaaggagcaggagcctttggcctctccaagggaagcctagtgggcctagacccctcc	804
Db	796	TKGAAGAGCAGGGAGC--TTTGCTCTCCAAGGGAAGCCTAGTGGCCCTAGACCCCTCC	739
QY	805	tccatggctagaagtgggcctgcaccatacatctgtgcgccccctctacccctcc	864
Db	738	TCCCATGGCTAGAAGT-GGGCTGCACCATACATCTGTGTCCGCCCTCTATACCCCTTCC	680
QY	865	cccacgtagggcactgtatggaccacgacgaggacagccatgggtcccgggcgccct	924
Db	679	CCCACTAGGGCACCTGTATGGACCAAGCACGGGACAGCATGGGTCCCGGGCGGCT	620
QY	925	tgtgctctggtaatgtttggtaccacaaacttggggccaaaaagggcagtgcctcaggact	984
Db	619	TGTGGCTCTGTTAATGTTTGGTACCAAACTTGGGGGCCAAAAGGCGAGTGCTCAGGACT	560
QY	985	ccctggccccctggtacctttccctgactcctggtgcctctcctcttgtccccccagaga	1044
Db	559	CCCTGGCCCCCTGGTACCTTTCCCTKACTCCTGGTGCCCTCTCCCTTTGTCCCCCAGAGA	500
QY	1045	gacatatgccccagagagagcaaatcgaagcgtgggaggcacccccattgctctcctcc	1104
Db	499	GACATATGCCCCCAGAGAGAGCAAAATCGAAGCGTGGAGGACACCCCATTTGCTCTCTCC	440
QY	1105	aggggcagaacatggggaggggactagatgggcaaggggcagcactgctgctctcct	1164
Db	439	AGGGCAGAACATGGGGAGGGGACTAGATGGGCAAGGGGCGACACTGCCCTGCTGCTCTCT	380
QY	1165	tcccctgtttacagcaataagcacgtcctcctccccacatccccacttccaggattgtggt	1224
Db	379	TCCCTCTGTTACAGCANAAGCACGCTCTCTCTCCCCAACTCCACTTCCAGGATTGTGGT	320
QY	1225	ttggattgaaaccaaagtttacaagtagacacccctggggggcggcagtggaacaaggat	1284
Db	319	TTGGATTGAAACCAAGTTTACAAGTAGACACCCCCCTGGGGGGCGGCAGTGRACAAGGAT	260
QY	1285	gcaaaggggtgggcattgggggtgccaggcaggaatgtacagactctatatctctatatat	1344
Db	259	GGCAAGGGGTGKGCAATGGGGTGCCAGGCAGGCATGTACAGACTCTATATCTCTATATAT	200
QY	1345	aatgtacagacagacagagtcccttccctctttaacccccctgacctttcttgactcccc	1404
Db	199	AATGAACAGACAGACAGAGTCCCTTCCMTCTTTAAACCCCTTGACCTTCTTGACTTCCCC	140
QY	1405	ttcagcttcagacaccccttccccaccaggctagggccccccacacctgggggacccctggc	1464
Db	139	TTCAGCTTCAGACCCCTTCCCCACCAGGCTTAGGCCCCCCCACACCTTGGGGGACCCCTGGC	80
QY	1465	ccctctttgtctctgtgaagacaggacctatgcaacgcacagacacttttgggagaccg	1524
Db	79	CCCTCTTTTCTCTGTGTGAAGACAGGACCTATGCAACGCACAGACACTTTTGGAGACCG	20
QY	1525	taaaacaacagcgcctccct	1543
Db	19	TAAWACAACAGCGCCCT	1

ORIGIN

Query Match		39.5%;	Score 687.8;	DB 10;	Length 702;
Best Local Similarity		99.6%;	Pred. No. 3e-84;		
Matches 700;		Conservative 0;	Mismatches 2;	Indels 1;	Gaps 1;
QY	555	tgaagatcaacgtgctggaagactttgaggggagagaacotcaggtgcccaagcttgaga	614		
Db	1				
QY	615	agagcatcagcgggaccagccccaaacgggaaacacctgccccctggccgtgggcacgcct	674		
Db	61				
QY	675	tcttcctcatgacgttcttggcctccttagctctgccccctccccctgggggggagagatg	734		
Db	121				
QY	735	ggcggggcttggaaagagcagggagcctttggcctctccaaagggaagcctagtgggcct	794		
Db	181				
QY	795	agacccctcctcccatggctagaaagtggggcctgcaccatacatctgtctcgccccctc	854		
Db	241				
QY	855	taccccttccccccacgtagggcactgtagtggaccaagcagcgggacagccatgggctc	914		
Db	301				
QY	915	cggcgggccttggctctggttaattgttggtaaccaacttgggggccccaaaagggcagt	974		
Db	361				
QY	975	gctcaggactccctggccccctggtacctttccctgactcctggtgccccctctcccttgtc	1034		
Db	421				
QY	1035	ccccagagagacatatgccccccagagagagcaaatcgaagcgtgggaggcacccccatt	1094		
Db	481				
QY	1095	gctctctccaggggcagaaacatggggaggggactagatgggcaaggggcagcactgcct	1154		
Db	541				
QY	1155	gctgcttctctccccctgtttacagcaataagcacgtcctcctccccactcccacttcca	1214		
Db	601				
QY	1215	ggattggtggttggattgaaacccaagtttacaagtagacacc	1257		
Db	660				
		GGATCGTGGTTTCGGATTGAACCAACCAAGTTTACAAGTAGACACCC	702		

RESULT	6
AL572018/c	
LOCUS	
DEFINITION	AL572018 LTI_NFL006_PL2 Homo sapiens CDNA clone CSODI026YI24 3 prime, mRNA sequence.
ACCESSION	AL572018
VERSION	AL572018.1 GI:12929883
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope

Genoscope - Centre National de Sequençage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. .750

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODI026Y124"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com>"

137 a 199 c 241 g 153 t 20 others

BASE COUNT
ORIGIN

Query Match	35.9%;	Score 625.4;	DB 9;	Length 750;
Best Local Similarity	94.3%;	Pred. No. 8.4e-76;		
Matches 706;	Conservative 16;	Mismatches 18;	Indels 9;	Gaps 7;

QY	802	tctcccatggtagaagtgggcctgtcaccatacatctgtgtccgccccctctaccct	861
Db	750		
QY	862	tccccacgtagggcactgtagtggaccaagcacgggagacgccatgggtcccgggcg	921
Db	690	:	
QY	922	cttgtggctctggttaattttgtaccaaacttgggggccccaa-aaggcagtgctcag	980
Db	630		
QY	981	gactccctggccccgttacctttccctgactcctgtgtgcctctctccctttgtccccca	1040
Db	570		
QY	1041	gagagacatatgccccagagagacaaatcgaaagcgtgggagggacccccattgctctc	1100
Db	511		
QY	1101	ctccaggggcagaaacatggggaggggactagatgggcaaggggcagcactgctgct	1160
Db	451		
QY	1161	tcttccccctgtttacagcaataagcacgtctctctccccactcccacttccaggattg	1220
Db	391		
QY	1221	tggtttgattgaaaccaagtttacaaagtagacacccccctggggggcgagtgacaa	1280
Db	331		
QY	1281	ggatgccaaaggggtggccattggggtgccaggcagcatgtacagactctatatct-cta	1339
Db	271		
QY	1340	tatataatgtacagacagacagatcccccttccctctttaacccccctgacctttcttgact	1399
Db	211	TATATAATGTACAGACAGACAGAGGCCCTTCCCTCTTTAACCCCCCTGACCTKWCYTGACT	152
QY	1400	tc-cccttcagcttcagaccccccttccccaccaggctaggccccca-cacctgggggacc	1457
Db	151		
QY	1458	ccctggccccctttttgtct---tctgtgaagacagggacctt-tgcaacgcacagacct	1513


```
Db 541 GAAGATMAACGTGCTGAAGACTTTGAGGGAGAGAACCCTCAGGTGCCCAAGCTTGAGAA 600
QY 616 gagcatcagcgggaccagcccccacaaacgggaacacactgccccctggcgtggcgcatt 675
Db 601 GAGCATCAGCGGGACAGCCCCAACAGCGGAACACCTGCCCTGGCCGTGGGCATCGCCTT 660
QY 676 ctctcctcatgacgtttcttgccc 697
Db 661 CTCCTCMGACGATATCTTGGC 682

RESULT 9
BF968137 826 bp mRNA linear EST 22-JAN-2001
LOCUS 602269124F1 NIH_MGC_84 Homo sapiens CDNA clone IMAGE:4357493 5',
DEFINITION mRNA sequence.
ACCESSION BF968137
VERSION BF968137.1 GI:12335352
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 826)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM993 row: p column: 06
High quality sequence stop: 769.
Location/Qualifiers
source 1..826
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4357493"
/clone_lib="NIH_MGC_84"
/tissue_type="adrenal cortex carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1:
NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT
primed. Average insert size 1.229 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 180 a 260 c 208 g 178 t
ORIGIN

Query Match 35.2%; Score 614.4; DB 10; Length 826;
Best Local Similarity 89.4%; Pred. No. 2.5e-74;
Matches 735; Conservative 0; Mismatches 66; Indels 21; Gaps 6;

QY 864 ccccccacgtaggcactgtagtggaccacgggacagccatgggtcccgccggcc 923
Db 1 CCCCCACGTAGGCACCTGTAGTGACCAAGCACGGGGACAGCCATGGGTCCCCGGCGGCC 60
QY 924 ttgtgctctggaatgtttgttaccacaaacttggggccacaaaggcagtgctcaggac 983
Db 61 TTGTGGCTCTGTAATGTTGGTACCAAACTTGGGGGCCAAAGGCGCAGTCTCAGGAC 120
QY 984 tccctggccccctggtaccttttccctgactcctggtgcccctctcctttgtccccccagag 1043
Db 121 TCCCTGGCCCCCTGGTACTTTCCCTGACTCCTGGTGCCTCTCCCTTTGTCTCCCCACAG 180
QY 1044 agacatatgccccagagagagcaaatcgaaagcgtggaggcacccccattgctcctc 1103
|||||
```

```
Db 181 AGACATATGCCCCCAGAGAGAGCAAAATCGAAGCGTGGGAGGACCCCCCATTTGCTCTC 240
QY 1104 caggggcagaacatgggaggggactagatgggcaaggggagcagcactgcctgctctcc 1163
Db 241 CAGGGGCAGAACATGGGAGGGGACTAGATGGCAAGGGGAGGAGGAGGAGGAGGAGGAG 300
QY 1164 ttccccctgtttacagcaataaagcagtcctcctccccccactccccacttccaggattgg 1223
Db 301 TTCCCCCTGTTTACAGCAATAAGCAGCTCCTCTCTCCCCACTTCCACTTCCAGGATTGG 360
QY 1224 tttagattgaaaccaaagtattacaagtagacacccccctggggggcggcagtggaagga 1283
Db 361 TTTGGATTGAACCAAGTTTACAACTAGACACCCCTGGGGGGGGGCGAGTGGACAAGGA 420
QY 1284 tgccaaaggggtgggcattgggtgcccaggcagggcatgtacagactctatatctatata 1343
Db 421 TGGCAAGGGGTGGGCATTGGGGTGCCAGGAGGCATGTACAGACTCTATATCTCTATA 480
QY 1344 taatgtacagacagacagagtcctccttcccttctttaaacccttgaccttcttgacctcc 1403
Db 481 TAATGTACAGACAGACAGAGTCCCTTCCCTCTTTAACCCCTTGACCTTTCTTGACTTCCC 540
QY 1404 cttcagcttcaga--ccccctcccccacaggttagggccccccacacacacacacacac 1461
Db 541 CTTCAAGCTTCAGACACCCCTTCCCAACAGGCTAGG-CCCCCACAACCTGGGGGACCCCT 599
QY 1462 gccccctctttgtcttctgtgaagacagacacacacacacacacacacacacacacacaga 1516
Db 600 GGGCCATCTTTGTCTTCTGTGAAGAACAGGACCTATGCTGACGCCACAGACCACTTTTG 659
QY 1517 ggagaccgtataaacaacagcgccccctccctccctccagccctgagccgggaaccatctccca 1576
Db 660 GAAGACCGTAAACAAACAAACAGCCCTCCCTTTACGCCCTGAGCCGGGAA----CATCTC 715
QY 1577 ggaccttgccctgctcaccctatgtgtgtccacacacacacacacacacacacacacacac 1636
Db 716 CAGGACTTGCCTGCTCACCCTATGTGTCCACTATCTCC-----GGGCTTTTCAAGTGC 769
QY 1637 ttgtgctgtgactttcactactctgtctcttagtctataaaaaa 1678
Db 770 TTGCTGTG--AATTCAATACCCTGCTCTTTGTTAAACAAAAA 808
```

```
RESULT 10
BF983120
LOCUS BF983120 804 bp mRNA linear EST 23-JAN-2001
DEFINITION 602305829F1 NIH_MGC_88 Homo sapiens CDNA clone IMAGE:4397263 5',
ACCESSION mRNA sequence.
VERSION BF983120
KEYWORDS BF983120.1 GI:12386017
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 804)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10097 row: i column: 08
High quality sequence stop: 766.
Location/Qualifiers
source 1..804
/organism="Homo sapiens"
```



```
|||||
Db 416 -GGAGGCACCCCATGCTCTCCTCCA-GGSCAGAACATGGGAGGGGACTAGATGGGCA 473
QY 1139 aggggcagcactgctgctgcttcccttccctgtttacagcaataagcacgctcctcc 1198
Db 474 AGGGCAGCACTGCTGCTGCTTCCCTGTTTACAGCAATAAGCACGCTCCTCCTCC 533
QY 1199 cccactcccactccaggattgtggttgattgaaaccaagtttacaagtagacaccc 1258
Db 534 CCAA-TCCCACCTCCAGGAT--GTGGTTGGATTGAACCAAGTTTACAAGTAGACACCCC 590
QY 1259 tggggggggcggcagtggaagaaggatgccaagggtgggcatgggggtgccaggcagga 1318
Db 591 TGGGGGGGGCGGCAGTGGACAAGGATGGCAAGGGGTGGGCANT-GGGTGCCAGGCAGGCA 649
QY 1319 tgtacagactctatatctctatatataatgtacagacagacagagtccttcccttctta 1378
Db 650 TGTACAGACTCTATATCTCTATATATAATGTACAGACAGACAGAGTCCTTCCCTCTTAA 709
QY 1379 accccctgaccttcttgacttc 1401
Db 710 CCCCAGGACTCTTGATTCCCTTC 732

RESULT 12
AL546822 AL546822 LTI_NFL006_PL2 902 bp mRNA linear EST 16-FEB-2001
LOCUS AL546822 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI026YI24 5
DEFINITION prime, mRNA sequence.
ACCESSION AL546822
VERSION AL546822.1 GI:12680311
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 902)
JOURNAL Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
JOURNAL Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
Location/Qualifiers
1..902
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI026YI24"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(df) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 161 a 297 c 278 g 156 t 10 others
ORIGIN

Query Match 31.7%; Score 553.2; DB 9; Length 902;
Best Local Similarity 91.1%; Pred. No. 4.9e-66;
Matches 749; Conservative 8; Mismatches 37; Indels 28; Gaps 15;

QY 105 ccaaccagcactgcggcagaggggtacaccgtgcaggtgaactgaa-----cga 156
Db 67 CCCACAGCCTGCGGCGACGCGGCTACACCGTGCAGGTGAACGTGCAGACGAGCTCAC 126
```

```
QY 157 ctatctggatattactgcccgcactacaacagctcgggggtggggcccgggcg---g 213
Db 127 TCGTGGATATGTTAGCTGCCCGCACTACAACAGCTCGGGGGTGGCCCCGGCGGAC 186
QY 214 accggggggcggaggcgggcaga-gcagtagc-gtctgtac-atggtgagcgcgaacg 270
Db 187 CGCGGGCGCGGAGGGCGGGGAGAGCGAGCGAGTACGCTGCTGTACGATGGTGAGCGCAACG 246
QY 271 ctaccgcacctgcaacgccagccagggttcaagcgtggaagtcgaacgcgcgcgc 330
Db 247 CTACCGCACCTGCAACGCCAGCGAGGGCTTCAAGCGTGGAGTGAACCGCGCGCACGC 306
QY 331 ccgcacagccccat-caagttctcggaagttccagcgctaca--gcgccttctct 387
Db 307 CCGGCACAGCCCCATGCAAGTCTCGGAGAAAGTTCCAGCGCTACAGCGCGCTCTCTCT 366
QY 388 gggctacagattccacgcgcgcagagtagtactactatctccacgcccactcaaacct 447
Db 367 GGGCTACGAGTTCCACGCCGCCACGAGTACTACTACTCTCCACGCCCTCACATCGCATCG 426
QY 448 gcactggaagtgtctgagatgaaggtgtctgtctgtcgcctccacatcgcaactcgg 507
Db 427 GCACTGGAAGTGTCTGAGATGAAGGTGTTCTGCTCTGCGCTCCACATCGCATCGCGG 486
QY 508 g-gagaagccggtcccactctcccagttccacatggcccggtggccttctctcat 566
Db 487 GCGAGAAGCCGGTCCCACACTCTCCCCAGTTTACCATTGCGGCGGCCCCAATGTGAAGATCAACG 546
QY 567 tgcgtgaagactttgaggggagagaacccctcaggtgccaaagcttgagaagagcat--cag 624
Db 547 TGCTGGAAGACTTTGAGGGAGRGAACCCCTCAGGTGCCCAAGCTTGAGAAGAGCATGCAGG 606
QY 625 cgggaccagccccaaacgggaacacacctgccccctggcgtggccttctctcat 684
Db 607 CGGGACCAGCCCCAAACGGGAACACCTGCCCCCTGGCGCTGGCATCGCTTCTCTCTCAT 666
QY 685 gacgttcttgccctctagctctgccccctccccct-ggggggggagagatggggcgggc 743
Db 667 GACGTTCTTGGCCTCCTAGTCTGTGCCCCCTCCCCYGGGGGGGAGAGATGGGGCGGGC 726
QY 744 ttggaaggagcaggagcctttggcctctccaaaggaaagccttagtggccttagacccctc 803
Db 727 TTGGAAGGAGCAGGAGAGCCTTTGGCCTCTCCAA-GGAARCTTAGTGGGCTAGA-CCCTC 784
QY 804 ctcccatggttagaagtggggctgacaccatacatctgtgtccgccccctctacccttc 863
Db 785 CTCCCATGGCTAGAAAGTGGGGSCTGCACCATACATCTGTGTCCGSCC---CTSTACCCYT 841
QY 864 cccccacgtagggcactgtagtggaccacagcacgggagacagc 905
Db 842 CCCCCACGTA-GGCACTGTAGTGA-CAAGCACGGGGGACABC 881
```

```
RESULT 13
BG499730
LOCUS BG499730 547 bp mRNA linear EST 27-MAR-2001
DEFINITION 602545144F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4667555 5',
mRNA sequence.
ACCESSION BG499730
VERSION BG499730.1 GI:13461247
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 547)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: CLONETECH Laboratories, Inc.
```


CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LCM1476 row: g column: 12
High quality sequence stop: 537.

FEATURES

source

1. .547
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4667555"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggccgctggcc); Site_2: SfiI (ggccattatggcc
); Double-stranded cDNA was prepared from cell line RNA.
5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGCGCCGACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."

BASE COUNT 112 a 171 c 146 g 118 t
ORIGIN

Query Match 30.4%; Score 530.2; DB 10; Length 547;
Best Local Similarity 99.3%; Pred. No. 7.3e-63;
Matches 543; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 898 gggacagccatgggtcccgggcgccctgtggtctgggtaattgtttgttaccacaaacttgg 957
Db 1 GGGACAGCCATGGTCCCGGGCGCCCTTGTGGCTCTGGTAATGTTTGGTACCAAACTTGG 60
QY 958 gggccaaaaagggcagtcagtcaggactccctggcccttggtaccttccctgactcctgg 1017
Db 61 GGGCCAAAAGGGCAGTGTCTCAGGACTCCCTGGCCCTGGTACCTTTCCCTGACTCCTGG 120
QY 1018 tgccctctccctttgtcccccagagagacatatgccccccagagagagacaaatcgaaagcg 1077
Db 121 TGCCCTCTCCCTTTGTCCCCCAGAGAGACATATGCCCCCAGAGAGAGCAAAATCGAAGCG 180
QY 1078 tgggaggcacccttcctctcctccaggggcagacacatggtgggggagggactagatgggc 1137
Db 181 TGGGAGGCACCCCATTTGCTCTCCTCCAGGGGCGAAGACATGGGGAGGGGACTAGATGGGC 240
QY 1138 aaggggcagcactgcctgcctcctcctccctgtttacagcaataagcactcctcctc 1197
Db 241 AAGGGCAGCAGTGGCTGCTGCTTCTTCCCTGTTTACAGCAATAAGCAGCTCCTCCTC 300
QY 1198 cccactcccactccaggattgtggtttgattgaaaccaaagtttacaagtagacaccc 1257
Db 301 CCCCACTCCCACTCCAGGATTGTGGTTTGAATTGAAACCAAGTTTACAAGTAGACACCC 360
QY 1258 ctggggggggcggcagtggaacaggatgccaaggggtgggcattggggtgccaggcagggc 1317
Db 361 CTGGGGGGGGCGAGTGGACAAAGGATGGCAAGGGGTGGGCATTGGGGTGGCAGGCAGGC 420
QY 1318 atgtacagactctatatctctatatataatgtacagacagacagagtcctctcctcttt 1377
Db 421 ATGTACAGACTCTATATCTATATATATGTACAGACAGACAGAGTCCCTTCCCTCTTT 480
QY 1378 aacccccctgactttcttgaacttcccccttcagcttcaga-ccccctccccaccaggctag 1436
Db 481 AACCCCTGACTTTCTTGACTTCCCTCTCAGCTTCAGACCCCTTCCCCACCAGGCTAA 540
QY 1437 gcccccc 1443
| | | | |

Db 541 GGCCCCC 547

RESULT 14

BF953157/c

LOCUS

DEFINITION

BF953157

ACCESSION

BF953157.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

BF953157 496 bp mRNA linear EST 22-JAN-2001
CM3-NN1187-111100-478-a06 NN1187 Homo sapiens cDNA, mRNA sequence.

BF953157.1 GI:12370432

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 496)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br.

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-NN1187-

111100-478-a06&t3=2000-11-11&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 19

High quality sequence stop: 495.

Location/Qualifiers

1. .496

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NN1187"

/dev_stage="Adult"

/note="Organ: nervous_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under

low stringency conditions."

BASE COUNT 94 a 144 c 162 g 96 t

ORIGIN

Query Match

Best Local Similarity 25.2%; Score 439.2; DB 10; Length 496;

Matches 466; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 443 aacctgcactggaagtgtctgaggatgaaggtgtctgtctgtgcctcc-acatcgca 501

Db 496 AACCTGCACTGGAAGTGTCTGAGGATGAAGGTGTTCTTCTGCTGGCCTCCACATCGCA 437

QY 502 ctccggggagaaagccgggtcccccactctccccccagttccaccatggggcccaatgtgaagat 561

Db 436 CTCCGGGAGAAAGCCGGTCCCCACTCTCCCCCAGTTCACCATGGGCCCCCAATATGAAGAT 377

QY 562 caacgtgctggaagacactttgagggagagaacccctcaggtgcccccaatgtgaagat 621

Db 376 CAACGTGCTGGAAGACTTTTGAGGGAGAGAAACCCCTCAGGTGCCCAAGCTTGAGAAGAGCAT 317

```
QY 622 cagcgggaccagcccccaacgggaacacacctgccccctggccgtgggcacatgcctcttctcct 681
|||||
Db 316 CAGCGGGACCAAGCCCAACGGAACACCTGCCCTGGCCGTGGGCATCGCCTTCTCCT 257
|||||
QY 682 catgacgttcttggcctccttagctctgcctccctccctcctgggggggagagatggggcgggg 741
|||||
Db 256 CATGACGTTCTTGGCCCTCCTAGCTCTGCCCCCTCCCTGGGGGGGGAGAGATGGGCGGG 197
|||||
QY 742 gcttgaaggagcaggagccttggcctctccaaagggaagcctagtggccttagacccc 801
|||||
Db 196 GCTTGAAGGAGCAGGAGCCCTTTGGCCCTCTCCAAAGGAAGCCTAGTGGGCGCTAGACCCC 137
|||||
QY 802 tctctccatggctagagtgggcctgcaccatacatctgtgtccgcctccctctacccct 861
|||||
Db 136 TCCTCCCATGGCTAGAGT-GGGCCTGCACCAATACATCTGTGTCCGCCCCCTCTACCCCT 78
|||||
QY 862 tccccccacgtagggcactgtagtggaccaagcacgggggacagccatgggtcccg 917
|||||
Db 77 TCCCCCACGTAGGGCACTGTAGTGGACCAAGCACGGGGACATACCGTGGTCCCGG 22
|||||

RESULT 15
BE138904/C
LOCUS
DEFINITION xw97e11.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2835980 3',
mRNA sequence.
ACCESSION BE138904
VERSION BE138904.1 GI:8601404
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 420)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 414.
FEATURES
Location/Qualifiers
source 1..420
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2835980"
/clone_lib="NCI_CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"

BASE COUNT 103 a 83 c 137 g 97 t
ORIGIN

Query Match 23.9%; Score 416.8; DB 9; Length 420;
Best Local Similarity 99.5%; Pred. No. 1.8e-47;
Matches 418; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1278 caaggatgccaaaggggtgggcattgggtgccaggcaggcatgtacagactctatatctc 1337
|||||
Db 420 CAAGGATGGCAAGGGGTGGGCATTGGGGTGCCAGGCAGGCATGTACAGACTCTATATCTC 361
|||||
QY 1338 tatatataatgtacagacagacagatcccttccctctttaacccctgaccttcttga 1397
|||||
```

```
Db 360 TATATATAATGTACAGACAGACAGAGTCCCTTCCCTCTTTAACCCCCCTGACCTTTCTTGA 301
QY 1398 ettcccccttcagcttcagaccccccttccccaccaggctagggccccccccacacacctgggggacc 1457
|||||
Db 300 CTTCCCCCTTCAGCTTCAGACCCCTTCCCCACCAGGCTAGGCCCCCCCACACACCTGGGGGACC 241
|||||
QY 1458 cccctggccccctcttctgtcttctgtgaagacagagacctatgcaacgcacagacaccttttg 1517
|||||
Db 240 CCCTGGCCCCCTCTTTTGTCTTCTGTGAAGACAGGACCTATGCAACGCACAGACACACTTTTG 181
|||||
QY 1518 gagaccgtaaaaaacaacagcgccccctcccttccagccctgagccgggaacccatctcccag 1577
|||||
Db 180 GAGACCGTAAACAACAACGCCCCCTCCCTTCCAGCCCTGAGCCGGGAACCATCTCCCCAG 121
|||||
QY 1578 gaccttgccccctgctcaccctatgtgggtccccaccctatcctcctgggacctttttcaagtgt 1637
|||||
Db 120 GACCTTGGCCCTGCTCACCCCTATGTGGTCCCACCTATCCTCCTGGGCCCTTTTCAAGTGCT 61
|||||
QY 1638 ttggctgtgacttttcatactctgctctttagtctaaaaaaaataaaactggagataaaaaata 1697
|||||
Db 60 TTGGCTGTGACTTTTCATACTCTGCTCTTAGTCTAAAAAAAATAAACTGGAGATAAAAAATA 1
```

Search completed: July 13, 2002, 09:10:59
Job time: 1839 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 13, 2002, 08:43:30 ; Search time 216.15 Seconds
(without alignments)
13844.931 Million cell updates/sec

Title: US-09-733-756-1
Perfect score: 1743
Sequence: 1 gctgctgctgctgctgctgc.....aaaaaaaaaaaaaaaaaaaaa 1743

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	949.2	54.5	1070	16	AAT03883 Eph transmembrane
2	917.8	52.7	1037	16	AAO85887 Human hek-L protei
3	447.8	25.7	715	21	AAA01637 Human colon cancer
4	252.2	14.5	257	21	AAA00162 Human colon cancer
5	205.8	11.8	358	21	AAC03424 Human secreted pro
6	149.2	8.6	642	20	AAX32767 Human LERK-6 polyp
7	149.2	8.6	783	17	AAT15009 Chicken elf-1 cDNA
8	149.2	8.6	783	19	AAV42927 cDNA encoding an a
9	139.6	8.0	1809	17	AAT15008 Mouse Elf-1 cDNA.

10	139.6	8.0	1809	19	AAV42926	cdna encoding a ma
11	135.4	7.8	314	17	AAT14010	LERK-6 exon. Homo
12	135.4	7.8	314	20	AAx32766	Exon sequence of h
13	128.8	7.4	555	17	AAT32700	Lerk-6 coding sequ
14	128.8	7.4	555	17	AAT14009	LERK-6 coding sequ
15	128.8	7.4	555	20	AAx32761	Murine LERK-6 poly
16	116.4	6.7	1438	13	AAO24595	B61 cDNA sequence.
17	116.4	6.7	1480	15	AAQ63770	B61 gene. Homo sa
18	116.4	6.7	1496	22	AAQ21387	Human cDNA sequenc
19	116.4	6.7	1496	22	AAC90574	Human PRO202 cDNA.
20	116.4	6.7	1496	22	AAC91576	Human PRO202 cDNA.
21	116.4	6.7	1527	21	AAC98823	Human pancreatic c
22	114.8	6.6	439	22	AAH81578	Human differential
23	102	5.9	687	17	AAT32699	Lerk-7 coding sequ
24	102	5.9	1176	23	AAS92024	DNA encoding novel
25	102	5.9	1574	23	AAS92779	DNA encoding novel
26	102	5.9	1728	17	AAT34292	HEK4 binding prote
27	102	5.9	1839	17	AAT18897	Human AL-1 cDNA.
28	101	5.8	636	16	AAQ85888	Human hek-L protei
29	73.2	4.2	3066	23	AAS75651	DNA encoding novel
30	61	3.5	374	22	AAI87907	Human polynucleoti
31	60.2	3.5	1631	22	AAH33111	Human polynucleoti
32	59.4	3.4	1542	22	AAD08420	Human colon cancer
33	59.2	3.4	491	22	AAI87701	Human secreted pro
34	59	3.4	836	22	AAI94700	Human polynucleoti
35	58.8	3.4	808	20	AAx89617	Human neuroblastom
36	58.8	3.4	808	22	AAS9248	Human secreted pro
37	58.8	3.4	808	24	ABA90917	Human polynucleoti
38	58.8	3.4	1898	22	AAF30056	Human cDNA encodin
39	58.8	3.4	1899	22	AAC91465	Human PRO306 cDNA.
40	58.4	3.4	1564	20	AAx22231	Human secreted pro
41	58.2	3.3	418	22	AAI90717	Human polynucleoti
42	58.2	3.3	612	22	AAH71471	Human polynucleoti
43	58	3.3	641	21	AAZ33356	Human cervical can
44	58	3.3	694	22	AAD06200	Human secreted pro
45	57.8	3.3	735	22	AAI02202	Human uteroglobin-
						Human reproductive

ALIGNMENTS

RESULT 1
AAT03883
ID AAT03883 standard; DNA; 1070 BP.
XX
AC AAT03883;
XX
DT 16-MAY-1996 (first entry)
XX
DE Eph transmembrane tyrosine kinase family ligand, Efl-2 encoding DNA.
XX
KW Efl-2; EHKL-L; Eph transmembrane tyrosine kinase family ligand;
KW neurological disorder; identification; diagnosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 46..750
FT /*tag= a
FT /product= Efl-2
XX
PN WO9527060-A2.
XX
PD 12-OCT-1995.
XX
PF 04-APR-1995; 95WO-US04208.
XX
PR 21-OCT-1994; 94US-0327423.
PR 04-APR-1994; 94US-0222075.
PR 12-APR-1994; 94US-0229402.
PR 01-SEP-1994; 94US-0299567.
XX
PA (REGE-) REGENERON PHARM INC.

Matches 512; Conservative 0; Mismatches 18; Indels 7; Gaps 5;

QY 1166 ccctgtttacagcaataaagcacgtctctccctcccccactccactccactccaggattgtggtt 1225
|||||
Db 52 ccctgtttacagcaataaagcacgtctctccctcccccactccca-ttccaggattgtggtt 110

QY 1226 tggattgaaaccaagttttacaagttagacacccctggggggcgaggcagtggaagagtg 1285
|||||
Db 111 tggattgaaaccaagttttacaagttagacacccctggggggcgaggcagtggaagagtg 170

QY 1286 ccaagggttgggcatgggtgcccaggcaggcagtgatgtacagactctatatctctatatata 1345
|||||
Db 171 gcaagggttgggcatgggtgcccaggcaggcagtgatgtacagactctatatctctatatata 230

QY 1346 atgtacagacagacagagtcctctccctcttttaacccctgacctttcttgacttccctt 1405
|||||
Db 231 atgtacagacagacagagtcctctccctcttttaacccctgacctttcttgacttccctt 290

QY 1406 tcagcttcagacccctcccccaccagcagtgaggcccccacacacccctggggaccctggcc 1465
|||||
Db 291 tcagcttcagacccctcccccaccagcagtgaggcccccacacacccctggggaccctggcc 350

QY 1466 cctctttgtcttctgtgaagacagagacctatgcaacgcacagacacttttggagaccgt 1525
|||||
Db 351 cctctttgtcttctgtgaagacagagacctatgcaacgcacagacacttttggagaccgt 410

QY 1526 aaaaacaac-agcgccccctccctccagccc-tgagccggggaaccatctccagaccctt 1583
|||||
Db 411 aaaaacaagcgccccctccctccagccc-tgagccggggaaccatctccagaccctt 470

QY 1584 gccctgtcacccctatgtgtgtccacacctatctctcctgggcccctttttcaa--gtgctttgg 1641
|||||
Db 471 gccctgtcacccctatgtgtgtccacacctatctctcctgggccccttttttnaagtgtttggg 530

QY 1642 ctgtgactttcactctgtctctt--agtcataaaaaataaaactggagataaaaaat 1696
|||||
Db 531 ctgtgactttcactctgtctctttagtctaaaaataaaactggagatnaaant 587

RESULT 4
AAA00162
ID AAA00162 standard; cDNA; 257 BP.
XX
AC AAA00162;
XX
DT 19-MAY-2000 (first entry)
XX
DE Human colon cancer cell line polynucleotide sequence SEQ ID NO:153.
XX
KW Human; colon cancer; tumour; diagnosis; gene expression product;
KW probe; detection; cancerous state; metastasis; identification;
KW breast cancer; oestrogen receptor-positive breast cancer; therapy;
KW oestrogen receptor-negative breast cancer; lung cancer; ss.
XX
OS Homo sapiens.
XX
PN WO9958675-A2.
XX
PD 18-NOV-1999.
XX
PF 13-MAY-1999; 99WO-US10602.
XX
PR 14-MAY-1998; 98US-0085426.
PR 15-MAY-1998; 98US-0085537.
PR 15-MAY-1998; 98US-0085696.
PR 21-OCT-1998; 98US-0105234.
PR 27-OCT-1998; 98US-0105877.
XX
(CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;

PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX
DR WPI; 2000-126369/11.
XX
PT Polynucleotide library used to determine cancerous states of mammalian
PT cells -
XX
PS Claim 1; Page 216; 1097pp; English.
XX
CC AAA0010 to AAA02716 represent polynucleotides isolated from cDNA
CC libraries constructed from human colon cancer cell lines. The present
CC invention also describes a method of detecting differentially expressed
CC genes correlated with a cancerous state of a mammalian cell, comprising
CC detecting at least one differentially expressed gene product in a test
CC sample derived from a cell suspected of being cancerous, where detection
CC of the differentially expressed gene product is correlated with a
CC cancerous state of the cell from which the test sample was derived.
CC The polynucleotides sequences can be used in a method for detecting
CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of
CC pre-metastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer.
XX
SQ Sequence 257 BP; 58 A; 74 C; 59 G; 66 T; 0 other;

Query Match 14.5%; Score 252.2; DB 21; Length 257;
Best Local Similarity 98.8%; Pred. NO. 1.6e-41;
Matches 254; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1166 ccctgtttacagcaataaagcacgtctctccctcccccactccactccaggattgtggtt 1225
|||||
Db 1 ccctgtttacagcaataaagcacgtctctccctcccccactccactccaggattgtggtt 60

QY 1226 tggattgaaaccaagttttacaagttagacacccctggggggcgaggcagtggaagagtg 1285
|||||
Db 61 tggattgaaaccaagttttacaagttagacacccctggggggcgaggcagtggaagagtg 120

QY 1286 ccaagggttgggcatgggtgcccaggcaggcagtgatgtacagactctatatctctatatata 1345
|||||
Db 121 gcaagggttgggcatgggtgcccaggcaggcagtgatgtacagactctatatctctatatata 180

QY 1346 atgtacagacagacagagtcctctccctcttttaacccctgacctttcttgacttccctt 1405
|||||
Db 181 atgtacagacagacagagtcctctccctcttttaacccctgacctttcttgacttccctt 240

QY 1406 tcagcttcagacccctt 1422
|||||
Db 241 ttagcttttagacccctt 257

RESULT 5
AAC03424
ID AAC03424 standard; cDNA; 358 BP.
XX
AC AAC03424;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 3422.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX

QY 401 cagccggccacgagtgactactacatct---ccacgcccaactcacaacctgcactggaag 457
Db 427 cggccggccacgagtgattactacatctctgccaagcctccaatgctgtggaccggccc 486
QY 458 tgtctgagatgaagggtgtctgtctgctgcgcctccacatgcactccgggg 509
Db 487 tgcctgcactgaagggtgtacgtgctggccgacccaacagagaccctgtacgag 538

RESULT 7
AAT15009
ID AAT15009 standard; cDNA; 783 BP.
XX
AC AAT15009;
XX
DT 02-JUL-1996 (first entry)
XX
DE Chicken elf-1 cDNA.
XX
KW Elf-1; EPH receptor ligand; dementia; tachycardia; gene therapy;
KW diagnosis; transgenic animal; ss; ds.
XX
OS Gallus sp.

Key Location/Qualifiers
5'UTR 1..85
CDS /*tag= a
sig_peptide 86..688
mat_peptide 86..148
3'UTR /*tag= b
polyA_signal 149..685
WO9609384-A1.
28-MAR-1996.
19-SEP-1995; 95WO-US11869.
27-FEB-1995; 95US-0393462.
19-SEP-1994; 94US-0308814.
(HARD) HARVARD COLLEGE.
Cheng H, Flanagan JG;
WPI; 1996-188446/19.
P-PSDB; AAR94767.
Murine and chicken EPH receptor ligand, Elf-1 - useful in diagnosis and treatment of disorders associated with the Elf-1 gene, e.g. dementia, tachycardia , etc.
Claim 36; Page 87-88; 107pp; English.

A chicken cDNA clone (AAT15009) codes for a novel EPH receptor ligand, Elf-1 (AAR94766), involved in the formation and maintenance of ordered spatial arrangements of differentiated tissue. It was obtd. from a day-3 chick embryo cDNA library using mouse Elf-1 cDNA (see AAT15008) as probe. The cDNA can be used for the prodn. of recombinant Elf-1, in the breeding of transgenic animals, for the design of diagnostic probes, and for gene (or antisense) therapy of cellular and tissue disorders.

Sequence 783 BP; 140 A; 270 C; 232 G; 141 T; 0 other;

Best Local Similarity 63.3%; Pred. No. 9.4e-21;
Matches 276; Conservative 0; Mismatches 133; Indels 27; Gaps 2;
QY 16 gctgctcgtgcccgctgccgctgctgccgctgctgcccacaggcccgagggcgctggg 75
Db 109 gctgctcgtgcccgctgctgctgctgctgctgctgctgctgctgctgctgctgct 168
QY 76 aaaccggcgcgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 135
Db 169 ggaccgctacgcccgtctattggaacgcagcaacccaggttccaccgctgctgctgct 228
QY 136 cgtgcaggtgaacgtgaacgtgaacgtgaacgtgaacgtgaacgtgaacgtgaacgt 195
Db 229 cgtgcaggtgaacgtgaacgtgaacgtgaacgtgaacgtgaacgtgaacgtgaacgt 279
QY 196 ggtggcccccggggcgccgg 255
Db 280 -----ggagccgctgcccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 324
QY 256 ggtgagccgcaacggctaccgcacacgtgcaacgcacggcc---agggcttcaagcgtgg 312
Db 325 ggtcaactacgaggggccacgcgtccctcgaccaccgcgcgcgcgcgcgcgcgcgcgcgc 384
QY 313 gtgcaacccggccgc 372
Db 385 gtgcaacccggccgc 444
QY 373 cagcgccttctctctggtggtggtggtggtggtggtggtggtggtggtggtggtggt 432
Db 445 cacccttctctctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 504
QY 433 gcccaactcacaacctg 448
Db 505 gtctcccccgaaactg 520

RESULT 8
AAV42927
ID AAV42927 standard; cDNA; 783 BP.
XX
AC AAV42927;
XX
DT 20-OCT-1998 (first entry)
XX
DE cDNA encoding an avian Elf-1 protein.
XX
KW Chicken; EPH receptor ligand; Elf-1; mek-4; sek-AP;
KW tyrosine kinase ligand; B61; LERK-2; proliferation; differentiation;
KW intracellular signalling; increased; survival; neuronal cell;
KW neuron survival; treatment; Alzheimer's; Parkinson's; lymphatic tumour;
KW artificial liver; cartilage; bone formation; ss.
XX
OS Gallus sp.

Key Location/Qualifiers
5'UTR 1..85
CDS /*tag= a
86..688
/*tag= b
/*product= Elf-1
sig_peptide 86..148
3'UTR /*tag= c
686..783
/*tag= d
US5795734-A.
18-AUG-1998.
31-MAY-1995; 95US-0455001.
31-MAY-1995; 95US-0455001.
19-SEP-1994; 94US-0308814.

Query Match 8.6%; Score 149.2; DB 17; Length 783;

Db 186 tcaagcgctggagtgcaaccgcccgcggcgccggggggcgctcaagttctcggaga 245
QY 360 agttccagcgctacagcgcttctctcttggtggtacaggttcacacgcccggccacgagtact 419
Db 246 agttccagctcttcacgccccttctctcctgggtctcgagttccggcccgcccgacgagtatt 305
QY 420 actacatct 428
Db 306 actacatct 314
RESULT 13
AAT32700
ID AAT32700 standard; cDNA; 555 BP.
XX
AC AAT32700;
XX
DT 28-NOV-1996 (first entry)
XX
DE Lerk-6 coding sequence.
XX
KW Lerk-6; hek; elk; cell surface receptor; culture; reagent;
KW neuron; disorder; injury; delivery agent; diagnostic; therapeutic;
KW Lerk-7; probe; cytokine; ss.
XX
OS Mus musculus.
XX
PN WO9617925-A1.
XX
PD 13-JUN-1996.
XX
PF 05-DEC-1995; 95WO-US15781.
XX
PR 01-MAR-1995; 95US-0396946.
PR 06-DEC-1994; 94US-0351025.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Cerretti DP;
XX
DR WPI; 1996-287171/29.
DR P-PSDB; AAW02587.
XX
XX New isolated human Lerk-7 cytokine - which binds to cell surface
PT receptors elk, hek and eck, useful for delivering agents to cells or
PT for treating neural disorders
XX
PS Example 1; Page 34; 49pp; English.
XX
CC A fragment of murine Lerk-6 DNA was isolated by PCR for use as a probe
CC for a human Lerk-6 homologue. The probe however led to the
CC identification of a human Lerk-7 gene (see AAT32699). The Lerk-7 gene and
CC protein can be used for studying the role of Lerk-7 in conjunction with
CC elk, hek and eck receptors. They can also be used for delivering
CC diagnostic or therapeutic agents to cells, e.g. cancer cells. The Lerk-7
CC proteins can also exhibit neuroprotective or neurotrophic properties and
CC can be used to treat neural tissue disorders.
XX
SQ Sequence 555 BP; 102 A; 191 C; 151 G; 111 T; 0 other;
Query Match 7.4%; Score 128.8; DB 17; Length 555;
Best Local Similarity 67.7%; Pred. No. 1.1e-16;
Matches 222; Conservative 0; Mismatches 82; Indels 24; Gaps 2;
QY 123 gagagggtacacccgtgcaggtgaacgtgaacgactatctggtatttactgcccgcact 182
Db 80 gcggcggtataccgtggaggtgagcatcaacgactacacctggtatctactgcccacact 139
QY 183 acaacagctcgggggtggcccccgggcgacccggcgccggagcgggcgagcagcagt 242
Db 140 a-----cggggcgcgctgcccccgctgacgcgcagcgcgt 178

QY 243 acgtgctgtacatggtgagccgcaacgggctaccgcacactgcaacgccagccagct 299
Db 179 acatcctgtacatggtgaatggtgagggccacgcctcctgtgaccacccgagcaggt 238
QY 300 tcaagcgctggagtgcaacccggccgcacgcccgcacagcccccatcaagtctcggaga 359
Db 239 tcaagcgctgggaatgcaacccggccgcagcgcgggggacccctcaagtctcagaga 298
QY 360 agttccagcgctacagcgcttctctctctggtggtacaggttcacacgcccggcagcagtagtact 419
Db 299 agttccaaactcttcaaccccttttccctgggttggagttccggtggccacgaatact 358
QY 420 actacatctcaccgcccactcaaacct 447
Db 359 actacatctcgcacacactcccaacct 386
RESULT 14
AAT14009
ID AAT14009 standard; cDNA; 555 BP.
XX
AC AAT14009;
XX
DT 10-OCT-1996 (first entry)
XX
DE LERK-6 coding sequence.
XX
KW LERK-6; hek; elk; cell surface receptors; culture; reagent;
KW neurons; disorder; injury; delivery agent; diagnostic; therapeutic;
KW ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 1..555
FT /*tag= a
FT /product= LERK-6
XX
PN WO9610911-A1.
XX
PD 18-APR-1996.
XX
PF 04-OCT-1995; 95WO-US12779.
XX
PR 03-OCT-1995; 95US-0538709.
PR 05-OCT-1994; 94US-0318393.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Cerretti DP;
XX
DR WPI; 1996-209575/21.
XX
PT Isolated DNA encoding cytokine designated LERK-6 which binds to hek
PT and elk cell surface receptors - useful for drug delivery and
PT screening procedures.
XX
PS Claim 3; Page 33; 44pp; English.
XX
CC The LERK-6 polypeptide encoded by this sequence can be used to
CC isolate cells expressing hek/elk cell surface receptors, or to
CC measure the biological activity of such receptors. The protein may
CC also be used as a delivery agent, taking diagnostic and therapeutic
CC agents to cells expressing such receptors. LERK-6 can also be used
CC as a tissue culture reagent to enhance the viability or prolong the
CC lifespan of the neurons. Neural tissue disorders and injuries may
CC be treated by contact with the polypeptide.
XX
SQ Sequence 555 BP; 102 A; 191 C; 151 G; 111 T; 0 other;
Query Match 7.4%; Score 128.8; DB 17; Length 555;
Best Local Similarity 67.7%; Pred. No. 1.1e-16;

	Matches	222;	Conservative	0;	Mismatches	82;	Indels	24;	Gaps	2;
QY	123	gagagggtacacccgtgaggtgaacgtgaacgactatctggataatttactgcccgcact	182							
Db	80	gcggcggtataccgtgaggtgagcatcaacgactacctgatatctactgccacact	139							
QY	183	acaacagctcgggggtgggccccggggcgggacccgggagcggggcagagcagt	242							
Db	140	a-----cggggcgcgctgccccgcgctgagcgcatggagcggt	178							
QY	243	acgtgctgtacatggtgagcgcgaacggctaccgcaactgcaacgccagccag---	299							
Db	179	acatcctgtacatggtgaatggtgagggccacgcctcctgtgaccaccggcagcgaggt	238							
QY	300	tcaagcgtgaggagtgaacacggccgcacgccccgcacagccccatcaagtctcggaga	359							
Db	239	tcaagcgtgggaatgaacacggccccgcagcgccccgggggagccccctcaagtctcagaga	298							
QY	360	agttccagcgtacagcgcccttctctctgtggtggtactagaggtccacgcccgcacgagtact	419							
Db	299	agttccaaactctccaccccttttccctgggctttgagttccggcctggccacgaatact	358							
QY	420	actacatctccacgcccactcacaacct	447							
Db	359	actacatctctgccacacctcccaacct	386							

RESULT 15
AAX32761
ID AAX32761 standard; cDNA; 555 BP.
XX
AC AAX32761;
XX
DT 24-JUN-1999 (first entry)
XX
DE Murine LERK-6 polypeptide encoding cDNA.
XX
KW LERK-6 polypeptide; hek receptor; elk receptor; human; murine; injury;
KW cell proliferation; neural growth; neural tissue; neurological disease;
KW neurodegenerative; excitotoxicity; ss.

CC effect on neural tissue. They can also be used as reagents for those
CC conducting quality assurance studies e.g. to monitor shelf life and
CC stability of elk protein under different conditions. The polypeptides can
CC also be used as carriers for delivering agents attached to cells bearing
CC the elk or hek cell surface receptor. The present sequence represents a
CC cDNA encoding a murine LERK-6 polypeptide.
XX
SQ Sequence 555 BP; 102 A; 191 C; 151 G; 111 T; 0 other;

Search completed: July 13, 2002, 09:53:34
Job time: 4204 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 13, 2002, 08:42:05 ; Search time 2227.85 Seconds
(without alignments)
16372.266 Million cell updates/sec

Title: US-09-733-756-1
Perfect score: 1743
Sequence: 1 gctgctgctgctgctgc.....aaaaaaaaaaaaaaaaaaaa 1743

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match, 100%
Listing first 45 summaries

Database : GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description

1	1678.8	96.3	1769	9	BC017722	BC017722 Homo sapi
2	1123.8	64.5	175826	2	AC021890	AC021890 Homo sapi
3	949.2	54.5	1070	6	AR004292	AR004292 Sequence
4	917.8	52.7	1037	6	AR001064	AR001064 Sequence
5	917.8	52.7	1037	6	AR080871	AR080871 Sequence
6	917.8	52.7	1037	6	AR164469	AR164469 Sequence
7	917.8	52.7	1037	6	I20809	I20809 Sequence 1
8	916	52.6	987	9	HSU14187	U14187 Human recep
9	704	40.4	758	9	HUMFL2	L37360 Homo sapien
c 10	643.2	36.9	289579	2	AC104632	AC104632 Mus muscu
11	642.2	36.8	326750	2	AC104327	AC104327 Mus muscu
12	253	14.5	1838	5	AB051678	AB051678 Danio rer
13	241.8	13.9	355	10	AY045577	AY045577 Rattus no
14	222	12.7	2535	10	U92885	U92885 Mus musculu
c 15	183.4	10.5	186	6	AX332137	AX332137 Sequence
16	159.2	9.1	264	5	AF209777	AF209777 Ctenophor
17	149.2	8.6	642	6	AR134683	AR134683 Sequence
18	149.2	8.6	642	6	AR152044	AR152044 Sequence
19	149.2	8.6	665	5	CHKELF1IG	L40932 Gallus gall
20	149.2	8.6	783	6	AR023765	AR023765 Sequence
21	147.6	8.5	1066	9	HS7292	AJ007292 Homo sapi
22	139.6	8.0	1540	10	MMU14941	U14941 Mus musculu
23	139.6	8.0	1594	10	MMU14752	U14752 Mus musculu
24	139.6	8.0	1809	6	AR023764	AR023764 Sequence
25	135.4	7.8	314	6	AR134682	AR134682 Sequence
26	135.4	7.8	314	6	AR152043	AR152043 Sequence
27	135.4	7.8	4686	9	HS6LK02	U92894 Homo sapien
28	135.4	7.8	36295	9	AC004258	AC004258 Homo sapi
29	130.2	7.5	799	5	GDRNARAGS	X90377 G.domesticu
30	128.8	7.4	555	6	ARI34677	ARI34677 Sequence
31	128.8	7.4	555	6	ARI52038	ARI52038 Sequence
32	125.2	7.2	1650	5	DTRKRAL1	Y09669 D.rerio mrn
33	121.8	7.0	307	10	AF131912	AF131912 Rattus no
c 34	121	6.9	162361	2	AC073818	AC073818 Mus muscu
35	119.8	6.9	651	5	XLU31204	U31204 Xenopus lae
36	119.4	6.9	492	5	XLU31205	U31205 Xenopus lea
37	119.4	6.9	2272	5	DTRKLELF1	Y09668 D.rerio mrn
38	119.2	6.8	188	10	MMU90666	U90666 Mus musculu
39	118.2	6.8	981	5	AF317286	AF317286 Gallus ga
40	116.8	6.7	649	10	MUSMB61B	D38146 Mouse B61 m
41	116.4	6.7	1480	6	AR048795	AR048795 Sequence
42	116.4	6.7	1480	6	AR103236	AR103236 Sequence
43	116.4	6.7	1480	6	I56902	I56902 Sequence 11
44	116.4	6.7	1480	6	I88172	I88172 Sequence 11
45	116.4	6.7	1480	9	HUMB61	M57730 Human B61 m

ALIGNMENTS

RESULT 1

BC017722

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

BC017722 Homo sapiens, ephrin-A3, clone MGC:21335 IMAGE:4397263, mRNA, linear PRI 06-DEC-2001
complete cds.
BC017722
MGC.
GI:17389356
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1769)
Straussberg, R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-ref@mail.nih.gov
Tissue Procurement: ATCC

Db 1645 CGGGAACCATCTCCAGGACCTTGCCCTGCTCACCCTATGTGGTCCACCTATCCTCCTG 1704

QY 1621 ggcctttttcaagtgttggctgtgacttttcatactctgtcttagtctataaaaaaata 1680
|||||

Db 1705 GGCCTTTTCAAGTGCTTTGGCTGTGACTTTCATACTCTGCTCTTAGTCTAAAAA 1764
|||||

QY 1681 aa 1682

Db 1765 AA 1766
||

RESULT 2

AC021890

LOCUS

DEFINITION Homo sapiens chromosome 3 clone RP11-498A2, WORKING DRAFT SEQUENCE,
22 unordered pieces.

AC021890

VERSION AC021890.12 GI:11128300

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 175826)

AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nickerson,S.,
Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojebokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Thomas,S., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Villalon,D., Vinson,R.,
Watlington,S., Williams,G., Williamson,A., Wlarczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.
and Gibbs,R.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 175826)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (22-JAN-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT On Nov 9, 2000 this sequence version replaced gi:9719697.
----- Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HMZE
Center clone name: RP11-498A2
----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 144734 bases at least Q40
Consensus quality: 162873 bases at least Q30
Consensus quality: 169424 bases at least Q20
Estimated insert size: 170770; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 3.3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 21031: contig of 21031 bp in length
* 21032 21131: gap of unknown length
* 21132 40475: contig of 19344 bp in length
* 40476 40575: gap of unknown length
* 40576 57617: contig of 17042 bp in length
* 57618 57717: gap of unknown length
* 57718 57708: contig of 17991 bp in length
* 57709 75808: gap of unknown length
* 75809 89018: contig of 13210 bp in length
* 89019 89118: gap of unknown length
* 89119 100348: contig of 11230 bp in length
* 100349 100448: gap of unknown length
* 100449 110704: contig of 10256 bp in length
* 110705 110804: gap of unknown length
* 110805 123156: contig of 12352 bp in length
* 123157 123257: gap of unknown length
* 123258 132467: contig of 9211 bp in length
* 132468 132567: gap of unknown length
* 132568 141461: contig of 8894 bp in length
* 141462 141561: gap of unknown length
* 141562 146576: contig of 5015 bp in length
* 146577 146676: gap of unknown length
* 146677 151201: contig of 4525 bp in length
* 151202 151301: gap of unknown length
* 151302 155722: contig of 4421 bp in length
* 155723 155822: gap of unknown length
* 155823 160148: contig of 4326 bp in length
* 160149 160248: gap of unknown length
* 160249 163024: contig of 2776 bp in length
* 163025 163124: gap of unknown length
* 163125 165339: contig of 2215 bp in length
* 165340 165439: gap of unknown length
* 165440 168113: contig of 2674 bp in length
* 168114 168213: gap of unknown length
* 168214 169867: contig of 1654 bp in length
* 169868 169967: gap of unknown length
* 169968 171580: contig of 1613 bp in length
* 171581 171680: gap of unknown length
* 171681 172882: contig of 1202 bp in length
* 172883 172982: gap of unknown length
* 172983 174431: contig of 1449 bp in length
* 174432 174531: gap of unknown length
* 174532 175826: contig of 1295 bp in length.

Location/Qualifiers
1. 175826

FEATURES
source

* 92955 96442: contig of 3488 bp in length
* 96443 96542: gap of unknown length
* 96543 98997: contig of 2455 bp in length
* 98998 99097: gap of unknown length
* 99098 102593: contig of 3496 bp in length
* 102594 102693: gap of unknown length
* 102694 105584: contig of 2891 bp in length
* 105585 105684: gap of unknown length
* 105685 108442: contig of 2758 bp in length
* 108443 108542: gap of unknown length
* 108543 111917: contig of 3375 bp in length
* 111918 112017: gap of unknown length
* 112018 116076: contig of 4059 bp in length
* 116077 116176: gap of unknown length
* 116177 119383: contig of 3207 bp in length
* 119384 119483: gap of unknown length
* 119484 123902: contig of 4419 bp in length
* 123903 124002: gap of unknown length
* 124003 127609: contig of 3607 bp in length
* 127610 127709: gap of unknown length
* 127710 131621: contig of 3912 bp in length
* 131622 131721: gap of unknown length
* 131722 135933: contig of 4212 bp in length
* 135934 136033: gap of unknown length
* 136034 140801: contig of 4768 bp in length
* 140802 140901: gap of unknown length
* 140902 144622: contig of 3721 bp in length
* 144623 144722: gap of unknown length
* 144723 148869: contig of 4147 bp in length
* 148870 148969: gap of unknown length
* 148970 153203: contig of 4234 bp in length
* 153204 153303: gap of unknown length
* 153304 157194: contig of 3891 bp in length
* 157195 157294: gap of unknown length
* 157295 162013: contig of 4719 bp in length
* 162014 162113: gap of unknown length
* 162114 167799: contig of 5686 bp in length
* 167800 167899: gap of unknown length
* 167900 173081: contig of 5182 bp in length
* 173082 173181: gap of unknown length
* 173182 180558: contig of 7377 bp in length
* 180559 180658: gap of unknown length
* 180659 186794: contig of 6136 bp in length
* 186795 186894: gap of unknown length
* 186895 192068: contig of 5174 bp in length
* 192069 192168: gap of unknown length
* 192169 198394: contig of 6226 bp in length
* 198395 198494: gap of unknown length
* 198495 203084: contig of 4590 bp in length
* 203085 203184: gap of unknown length
* 203185 208688: contig of 5504 bp in length
* 208689 208788: gap of unknown length
* 208789 214457: contig of 5669 bp in length
* 214458 214557: gap of unknown length
* 214558 219456: contig of 4899 bp in length
* 219457 219556: gap of unknown length
* 219557 226019: contig of 6463 bp in length
* 226020 226119: gap of unknown length
* 226120 234001: contig of 7882 bp in length
* 234002 234101: gap of unknown length
* 234102 241499: contig of 7398 bp in length
* 241500 241599: gap of unknown length
* 241600 249134: contig of 7535 bp in length
* 249135 249234: gap of unknown length
* 249235 259781: contig of 10547 bp in length
* 259782 259881: gap of unknown length
* 259882 267444: contig of 7563 bp in length
* 267445 267544: gap of unknown length
* 267545 279432: contig of 11888 bp in length
* 279433 279532: gap of unknown length
* 279533 289579: contig of 10047 bp in length.

FEATURES
source

Location/Qualifiers
1. .289579

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="rp23-368d24"
/clone_lib="RPCI mouse BAC library 23"
BASE COUNT 70974 a 71815 c 70765 g 68626 t 7399 others
ORIGIN

Query Match 36.9%; Score 643.2; DB 2; Length 289579;
Best Local Similarity 81.3%; Pred. No. 1.4e-114;
Matches 920; Conservative 0; Mismatches 163; Indels 49; Gaps 13;

QY 572 gaagactttgaggagagagaaccctcaggtgcccagcttgagaagagcatcagcgggacc 631
|||||
Db 184805 GAAGACTTTGAGGGAGAGAAATCCCCAGGTGCCCAAGCTTGAGAAAGAGCATCAGTGGGACC 184746

QY 632 agcccaaacgggaacacacctgccccctggccgtgggcatcgcttcttcctcatgacgttc 691
|||||
Db 184745 AGCCCCAAGCGGGAACACCTGCCCTCTGGCCGTGGGCATCGCCTTCTTCTCATGACGCTC 184686

QY 692 ttggcctcctagctctgccccctccccctggggggggagagatggggcgggcttggaaag 751
|||||
Db 184685 TTGGCCTCCTAGCTGTGCCCTTCTTATGACAGAGAGAGAGG--TGGGAAGGGCTGAGAAGG 184628

QY 752 agcaggaggacctttggcctctccaaagggaagcctagtgggcctagaccctcctcccatg 811
|||||
Db 184627 AGCAGGGAGCTT-----GCTGTGGGGCCTACATCCTTCTTCCCATG 184587

QY 812 gctagaagtggggcctgcaccatacatctgtgtccgccccctct-----accccttccc 865
|||
Db 184586 GTTGGGAGCGGGGTCTGCACTGTACATCTCTTGGCCTGCCCCTCTTGTCCCAACACTC 184527

QY 866 ccacacgtaggcactgtagtggaccaagcacggggacagccatgggtcccggcgccctt 925
|||
Db 184526 TCTAGGACAGSCACCGTAGTGGATCAGSCACAGGACAGCCACAGGTCCCAGGTGGCC-T 184468

QY 926 gtgctctgtgtaattgttgggtaccaaaacttggggggcccaaaaaggcagctgctcaggactc 985
|||||
Db 184467 GTGGCTTTGGTAATGTTGGTGTACCAAACTGGGGGCTATAAA-GGCAGTGCTCAGGACTC 184409

QY 986 cctggccccctggtacacctttccctgactcctggtgctcctctccttctgtccccccagagag 1045
|||||
Db 184408 CCTGGCCCCCTGGTACCTTTCCCTGACCCCTTGGTGCCCTCCCTCTTGTGTCCCCCAAGAGGC 184349

QY 1046 acatatgccccccagagagagcaaatcgaagcgtggaggacccccattgctctctccca 1105
|||
Db 184348 AAATATG-CCCCAGAGAGAAACAAA-CGAAGCATGGGAGGTGCCCTCTCTCTCTCT-CT 184292

QY 1106 ggggcagaacatggggagggggactagatgggcaagggggagcactgctgcttccctt 1165
|||||
Db 184291 GGGGCAGAACATGGGAGGGGACTAGGTGGGTGAGGGGTGAGGCTCAGGCTGC-CCCTC 184233

QY 1166 cccctgtttacagcaataaagcacgtcctcctccccccactccacttccaggattgtggtt 1225
|||||
Db 184232 CCCCTGTTTACAGCAATAAGCATGTCTCTCCCTCCACTCCACATAGAGGACTGTGGTT 184173

QY 1226 tggattgaaaccaaagtttacagtagacacccccctggggggggcagtggaacaaaggatg 1285
|||||
Db 184172 TGGATTGAATCCAAGTTTACAAATAGACACCCCTGGGGGGGAGCAGGAGCAG-----TG 184121

QY 1286 ccaaggggtgggcattggggtggccaggcaggcatgtacagactctatatctctatatata 1345
|||||
Db 184120 GACAGGGGTGGGCATCGGGGTGCCAGACAGGCGATGTACAGACTCTATATCTCTATATATA 184061

QY 1346 atgtacagacagacagagtccttccctctttaaacccttcccttcttctgacttccctt 1405
|||||
Db 184060 ATGTACAGACAGACAGAGTCCCTTCCCTTCTTAACTTGACCTTCTCTCGACTTCCCTT 184001

QY 1406 tcagcttcagaccccccttccccaccaggctaggccccccccacacacctgggggacccccctggcc 1465
|||
Db 184000 CCAGACTCAGACCCCTTCCCTTACCAGGTTAGGCCCC-----CCTTGGGACCCCCCTGGCC 183945


```
* 159448 159547: gap of unknown length
* 159548 163961: contig of 4414 bp in length
* 163962 164061: gap of unknown length
* 164062 169232: contig of 5171 bp in length
* 169233 169332: gap of unknown length
* 169333 173172: contig of 3840 bp in length
* 173173 173272: gap of unknown length
* 173273 178980: contig of 5708 bp in length
* 178981 179080: gap of unknown length
* 179081 188181: contig of 9101 bp in length
* 188182 188281: gap of unknown length
* 188282 192848: contig of 4567 bp in length
* 192849 192948: gap of unknown length
* 192949 198249: contig of 5301 bp in length
* 198250 198349: gap of unknown length
* 198350 206031: contig of 7682 bp in length
* 206032 206131: gap of unknown length
* 206132 210282: contig of 4151 bp in length
* 210283 210382: gap of unknown length
* 210383 218415: contig of 8033 bp in length
* 218416 218515: gap of unknown length
* 218516 226120: contig of 7605 bp in length
* 226121 226220: gap of unknown length
* 226221 236580: contig of 10360 bp in length
* 236581 245273: contig of 8593 bp in length
* 245274 245373: gap of unknown length
* 245374 256115: contig of 10742 bp in length
* 256116 256215: gap of unknown length
* 256216 264078: contig of 7863 bp in length
* 264079 264178: gap of unknown length
* 264179 273637: contig of 9459 bp in length
* 273638 273737: gap of unknown length
* 273738 295230: contig of 21493 bp in length
* 295231 295330: gap of unknown length
* 295331 313298: contig of 17968 bp in length
* 313299 326750: gap of unknown length
* 313399 326750: contig of 13352 bp in length.

FEATURES             Location/Qualifiers
     source            1..326750
                        /organism="Mus musculus"
                        /strain="C57BL/6J"
                        /db_xref="taxon:10090"
                        /clone="rp23-295a4"
                        /clone_lib="RPCI mouse BAC library 23"
BASE COUNT      79344 a 80488 c 78974 g 81219 t 6725 others
ORIGIN
```

Query Match 36.8%; Score 642.2; DB 2; Length 326750;
Best Local Similarity 81.2%; Pred. No. 2.3e-114;
Matches 919; Conservative 0; Mismatches 164; Indels 49; Gaps 13;

```
QY 572 gaagactttgagggagagaacccctcaggtgccccagcttgagaagagcatcagcgggacc 631
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195361 GAAGACTTTGAGGGAGAGAAATCCCGAGGTGCCCAAGCTTGAGAAGAGCATCAGTGGGACC 195420

QY 632 agcccaaacgggaacacccctgcccctggccgtgggcacgccttcttccatgacgttc 691
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195421 AGCCCAAGCGGGAACACCTGCCTCTGCGCGTGGGCATCGCCTTCTTCTCATGACGCTC 195480

QY 692 ttggcctcctagctctgccccctccccctgggggggagagatggggcgggcttggaagg 751
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195481 TTGGCCTCTAGCTCTGCCCTTCTCTATGACAGAGAGAGG--TGGGAAGGCTGAGAAGG 195538

QY 752 agcagggagcctttggcctctccaagggaagcctagtgggcctagacccctcctccatg 811
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195539 AGCAGGGAGCTT-----GCTGTGGGCGCTACATCCTTCTTCTCCATG 195579

QY 812 gctagaagtgggcccgtgcaccatacatctgtgtccgccccctct-----acccttccc 865
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 195580 GTTGGGACGGGGTCTGCACTGTACATCTCTCTGGGCGCTGCCCTCTTTGCCCCACACTC 195639
```

```
QY 866 cccacgtagggcactgtagtggaccaagcacgggggacagccatgggtccccgggcgctt 925
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195640 TTAGGACAGGCACCGTAGTGGATCAGGCACAGGCACAGCCACAGGTCCCAGGTGGCC-T 195698

QY 926 gtggctctggtaatgtttggtaccaaaacttggggggccaaaaggcagtgctcaggactc 985
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195699 GTGGCTTTGGTAATGTTGGTACCAAAACCTGGGGGCTATAAA-GGCAGTGCTCAGGACTC 195757

QY 986 cctggccccctggtacaccttccctgactcctggtgcccccttcccttcttgcctccagagag 1045
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195758 CCTGGCCCCCTGGTACCTTTCCTGACCCCTTGGTGCCCTCCCTCTTTGTCCCCCAAGAGGC 195817

QY 1046 acatatgccccccagagagagcaaatcgaaagctgggagggcaccccccatgtctctcccca 1105
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195818 AAATATG-CCCCAGAGAGAAACAAA-CGAAGCATGGGAGGTGCCCCCTGTCTCTCTCT-CT 195874

QY 1106 ggggcagaacatggggagggggactagatgggcaagggggcagcactgctgcttccct 1165
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195875 GGGGCAGAACATGGGGAGGGGACTAGTGGGTGAGGGGTGGAGCCTCAGGNTGC-CCCTC 195933

QY 1166 cccctgtttacagcaataagcacgtctcctctccctccacactccacattccaggattgtggtt 1225
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195934 CCCCTGTTTACAGCAATAAGCATGTCTCTCCCTCCACTCCACATAGAGGACTGTGGTT 195993

QY 1226 tggattgaaaccaaagtttacaagtacacacccctggggggcggcagtggaagagatg 1285
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195994 TGGATTGAATCCAAAGTTTACAAATAGACACCCCTGGGGGGAGCAGCAG-----TG 196045

QY 1286 ccaaggggtgggcattgggtgccaggcagggcatgtacagactctatatctctatatata 1345
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196046 GACAGGGGTGGGCATCGGGTGCACAGAGGCATGTACAGACTCTATATCTCTATATATA 196105

QY 1346 atgtacagacagacagatcccttccctctttaaacccttgaaccttcttgacttccct 1405
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196106 ATGTACAGACAGACAGAGTCCCTTCCCTTCTTAACCTTGACCTTCTCGACTTCCCTT 196165

QY 1406 tcagcttcagaccccccttccaccaggtgtagccccccacacacctggggacccccctggcc 1465
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196166 CCAGACTCAGACCCCTTCCCTACCAGGTAGGCCCC-----CCTGGGACCCCTGGCC 196221

QY 1466 cctcttctgtctgtgaagacagggacctatgcaacgcacagacacttttggagaccgt 1525
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196222 CCTCTTTGTCTCTGTGAAGACAGGACCTATGCAACGCACAGACACACTTTTGGAGACCGT 196281

QY 1526 aaaaacacagcggccccctccctccagccctgagccggggaacctctcccaggaccttgc 1585
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196282 --AACAAACAATGCCCCCTCCCTTCCAGCCCTGAGCCGGAACCAACTCCCAGGACCTGC 196339

QY 1586 cctgctcacacctatggtgtccacctatcctcctctgggccttttcaagtgtcttgctgt 1645
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196340 CCTGCCACCCCTAAGTGGTCC--CCACCCATCCTGGGCTTTTCAAGTGTCTTGGCTGT 196397

QY 1646 gacttcatactctgctcttagtctaaaaaaaataaaactggagataaaaaa 1697
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196398 GACTTTCATACTCTGCTCTTAGTCTAAAAAAAATAAACTGGAGATAAAAAATA 196449
```

RESULT 12
AB051678
LOCUS
DEFINITION
AB051678
VERSION
AB051678.1 GI:14196230
KEYWORDS
SOURCE
ORGANISM
Danio rerio
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 1838)
Hirate,Y., Mieda,M., Harada,T., Yamasu,K. and Okamoto,H.
Identification of ephrin-A3 and novel genes specific to the
midbrain-MHB in embryonic zebrafish by ordered differential display

AB051678 1838 bp mRNA linear VRT 20-SEP-2001
Danio rerio mRNA for ephrin-A3, complete cds.
AB051678
AB051678.1 GI:14196230

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

```

Run on:      July 13, 2002, 08:42:45 ; Search time 53.99 Seconds
              (without alignments)
              7929.972 Million cell updates/sec

```

Title: US-09-733-756-1
Perfect score: 1743
Sequence: 1 aactactactacta

C# / H

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```

database :
Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/pctus_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	949.2	54.5	1070	1	US-08-299-567-8	Sequence 8, Appli
2	917.8	52.7	1037	1	US-08-299-567-8	Sequence 1, Appli
3	917.8	52.7	1037	1	US-08-299-567-8	Sequence 1, Appli
4	917.8	52.7	1037	2	US-09-453-943-1	Sequence 1, Appli
5	917.8	52.7	1037	2	US-09-453-943-1	Sequence 1, Appli
6	149.2	8.6	642	1	US-09-358-734-1	Sequence 1, Appli
7	149.2	8.6	642	1	US-09-358-734-1	Sequence 1, Appli
8	149.2	8.6	642	2	US-09-609-324A-9	Sequence 9, Appli
9	149.2	8.6	642	4	US-08-920-440B-9	Sequence 9, Appli
10	149.2	8.6	642	4	US-09-173-492-9	Sequence 9, Appli
11	149.2	8.6	783	1	US-09-173-133-9	Sequence 9, Appli
12	139.6	8.0	1615	5	US-08-455-001-3	Sequence 3, Appli
13	139.6	8.0	1615	4	PCT-US95-11869-3	Sequence 3, Appli
14	139.6	8.0	1809	1	US-08-308-814-1	Sequence 1, Appli
15	135.4	7.8	1809	5	US-08-455-001-1	Sequence 1, Appli
16	135.4	7.8	314	1	PCT-US95-11869-1	Sequence 1, Appli
17	135.4	7.8	314	2	US-09-609-324A-7	Sequence 7, Appli
18	135.4	7.8	314	4	US-08-920-440B-7	Sequence 7, Appli
19	135.4	7.8	314	4	US-09-173-492-7	Sequence 7, Appli
20	128.8	7.4	314	4	US-09-173-133-7	Sequence 7, Appli
21	128.8	7.4	555	1	US-09-165-533-7	Sequence 7, Appli
22	128.8	7.4	555	2	US-09-609-324A-1	Sequence 7, Appli
23	128.8	7.4	555	2	US-08-920-440B-1	Sequence 1, Appli
24	128.8	7.4	555	4	US-09-173-492-1	Sequence 1, Appli
25	128.8	7.4	555	4	US-09-173-133-1	Sequence 1, Appli
26	128.8	7.4	555	4	US-09-165-533-1	Sequence 1, Appli
27	116.4	6.7	1480	5	PCT-US95-12779-1	Sequence 1, Appli
				5	PCT-US95-15781-1	Sequence 1, Appli
				1	US-08-448-736-11	Sequence 11, Appli

28	116.4	6.7	1480	1	US-08-452-779-11	Sequence 11, Appl
29	116.4	6.7	1480	1	US-08-445-065-11	Sequence 11, Appl
30	116.4	6.7	1480	3	US-08-959-524-11	Sequence 11, Appl
31	116.4	6.7	1498	1	US-08-321-162-1	Sequence 1, Appl
32	116.4	6.7	1498	1	US-08-441-216-1	Sequence 1, Appl
33	102	5.9	687	5	PCT-US95-15781-4	Sequence 1, Appl
34	102	5.9	1728	3	US-08-379-802-1	Sequence 4, Appl
35	102	5.9	1728	3	US-09-048-129-1	Sequence 1, Appl
36	102	5.9	1728	4	US-09-048-079-1	Sequence 1, Appl
37	102	5.9	1839	1	US-08-442-248-3	Sequence 1, Appl
38	102	5.9	1839	1	US-08-440-815-3	Sequence 3, Appl
39	102	5.9	1839	4	US-08-486-449-3	Sequence 3, Appl
40	101	5.8	636	1	US-08-240-124-3	Sequence 3, Appl
41	101	5.8	636	1	US-08-453-943-3	Sequence 3, Appl
42	101	5.8	636	2	US-09-057-121-3	Sequence 3, Appl
43	101	5.8	636	4	US-09-358-734-3	Sequence 3, Appl
44	60.8	3.5	7218	1	US-08-232-463-14	Sequence 14, Appl
45	58.4	3.4	1069	4	US-09-372-422A-7	Sequence 7, Appl

ALIGNMENTS

```

RESULT 1
US-08-299-567-8
; Sequence 8, Application US/082999567
; Patent No. 5747033
; GENERAL INFORMATION:
; APPLICANT: Davis, et al.
; TITLE OF INVENTION: METHOD OF ENHANCING THE BIOLOGICAL
; TITLE OF INVENTION: ACTIVITY OF EPH FAMILY LIGANDS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10591-6707
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,567
; FILING DATE: 01-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kempler, Gail M.
; REGISTRATION NUMBER: 32,143
; REFERENCE/DOCKET NUMBER: REG 290
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1070 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-299-567-8

```

Query Match	54.5%	Score 949.2;	DB 1;	Length 1070;
Best Local Similarity	97.0%	Pred. No. 9e-186;		
Matches 976; Conservative	0;	Mismatches 18;	Indels 12;	Gaps 1;

Qy 1 gctgctgctgctgctgctgctgctgcccgtgccgctgctgccgctgctggcccaaggcc 60
 |||||
Db 60 GCTGCTGCTGCTGCTGCTGCTGCTGCCCGTGCCCGCTGCTGCCGCTGCTGGCCCCAAGGGCC 119

Db 229 CGTGGAGGTGAGCATCAATGACTACCTGGACATCTACTGCCCTCACTACGA----- 279
QY 196 ggtggccccgggacccggggcccgaggcggggcagagcagtagctgctgtacat 255
Db 280 -----GGAGCGCTGCCCGCGGAGCGCATGGAGCGCTACGTCCTCTACAT 324
QY 256 ggtgagcgcgaacggctacgcacacgtgcaacgcagcc---agggcttcaagcgctgga 312
Db 325 GGTCAACTACGAGGGCCACGCGTCTCGGACCAACCGGAGGCTTCAAAACGTTGGGA 384
QY 313 gtgcaaccggccgcacgccccgcacagcccccatcaagtctcggagaagttccagcgcta 372
Db 385 GTGCAACCGGCGGACTCCCCAGCGGACCCCTCAAGTTCTCAGAGAAGTTCCAGCTCTT 444
QY 373 cagcgcttctctctggtgctacgagttccacgcggcccgccacgagtagtactactatctccac 432
Db 445 CACCCCTTCTCTTTGGCTTCGAGTTCCGTTCCGTCGCCGCCACGAGTATTACTATCTCTGC 504
QY 433 gcccaactcacaaacctg 448
Db 505 GTCTCCCCGGAACGTG 520

RESULT 11
PCT-US95-11869-3
; Sequence 3, Application PC/TUS9511869
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: EPH Receptor Ligands, and Uses Related
; TITLE OF INVENTION: Thereto
; NUMBER OF SEQUENCES: 5
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/11869
; FILING DATE: 19-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-011CPPC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 783 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 86..685
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..85
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 686..783
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 86..148
PCT-US95-11869-3

Query Match 8.6%; Score 149.2; DB 5; Length 783;
Best Local Similarity 63.3%; Pred. No. 4.2e-22;
Matches 276; Conservative 0; Mismatches 133; Indels 27; Gaps 2;
QY 16 gctgctggtgccccgtgccccgtgctgccccgtgccccgaagggccccggcgctggg 75

Db 109 GCTGCTCGCGCGATCGTGGCGTCTGCGTGTGGAGCGACGACCCCGGAAAGTATCTC 168
QY 76 aaaccggcatgcggtgtactggaacagctccaaccagcacctgcgggcagagggctacac 135
Db 169 GGACCGCTACGCCGTCTATTGGAACCGCAGCAACCCAGGTTCCACCGGGGATTACAC 228
QY 136 cgtgcaggtgaacgtgaacgactatctggatatattactgcccgcactacaacagctcggg 195
Db 229 CGTGGAGGTGAGCATCAATGACTACCTGGACATCTACTGCCCTCACTACGA----- 279
QY 196 ggtggccccgggacccggggcccgaggcggggcagagcagtagctgctgtacat 255
Db 280 -----GGAGCGCTGCCCGCGGAGCGCATGGAGCGCTACGTCCTCTACAT 324
QY 256 ggtgagcgcgaacggctacgcacacgtgcaacgcagcc---agggcttcaagcgctgga 312
Db 325 GGTCAACTACGAGGGCCACGCGTCTCGGACCAACCGGAGGCTTCAAAACGTTGGGA 384
QY 313 gtgcaaccggccgcacgccccgcacagcccccatcaagtctcggagaagttccagcgcta 372
Db 385 GTGCAACCGGCGGACTCCCCAGCGGACCCCTCAAGTTCTCAGAGAAGTTCCAGCTCTT 444
QY 373 cagcgcttctctctggtgctacgagttccacgcggcccgccacgagtagtactactatctccac 432
Db 445 CACCCCTTCTCTTTGGCTTCGAGTTCCGTTCCGTCGCCGCCACGAGTATTACTATCTCTGC 504
QY 433 gcccaactcacaaacctg 448
Db 505 GTCTCCCCGGAACGTG 520

RESULT 12
US-08-308-814-1
; Sequence 1, Application US/08308814
; Patent No. 6268476
; GENERAL INFORMATION:
; APPLICANT: Flanagan, John G.
; APPLICANT: Cheng, Hwai-Jong
; TITLE OF INVENTION: EPH Receptor Ligands, and Uses Related
; TITLE OF INVENTION: Thereto
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(txt)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/308,814
; FILING DATE: 19-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1615 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:

